Vol. 50 JUNE 1959 No. 3

The Journal of Educational Psychology

CONTENTS

Goal-Setting Behavior as a Function of Self-Acceptance, Over- and Underachievement, and Related Personality Variables 93
JAMES V. MITCHELL, JR.

Predicting Leadership Ratings From High School Activities 105
JOHN D. KRUMBOLTZ, RAYMOND E. CHRISTAL, AND JOE H. WARD, JR.

Effects of Social Recognition Upon the Educational Motivation of Talented Youth 111
DONALD L. THISTLETHWAITE

Intercorrelations Among Various Intelligence, Achievement, and Social Class Scores 117
LOTUS M. KNIEF AND JAMES B. STROUD

Some Behaviors of Elementary School Children Related to Classroom Activities and Subject Areas 121
SURANG KOWATRAKUL

Refusals and Illegibilities in the Spelling Errors of Maladjusted Children 129
HARRY H. L. KITANO

Personality and Independent Study 132
KATHRYN KOENIG AND W. J. McKEACHIE

AMERICAN PSYCHOLOGICAL ASSOCIATION

The Journal of Educational Psychology

RAYMOND G. KUHLEN, Editor
Syracuse University

Advisory Editors

LEE J. CRONBACH
University of Illinois
CHESTER W. HARRIS
University of Wisconsin
ERNEST R. HILGARD
Stanford University
NICHOLAS HOBBS
George Peabody College for Teachers
HAROLD E. JONES
University of California
IRVING LORGE
Teachers College, Columbia University
DOROTHEA McCARTHY
Fordham University
T. ERNEST NEWLAND
University of Illinois
H. H. REMMERS
Purdue University
JULIAN C. STANLEY
University of Wisconsin
J. B. STROUD
State University of Iowa
DONALD E. SUPER
Teachers College, Columbia University
PERCIVAL M. SYMONDS
Teachers College, Columbia University
GEORGE G. THOMPSON
Syracuse University

The Journal of Educational Psychology publishes original investigations of problems of learning, teaching, and the psychological development of the individual. It contains articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, personality, social relations, and vocational orientation, particularly (but not exclusively) as related to school adjustment or the impact of the school program upon development and adjustment; emotion, motivation, and character; mental development; and methods. The last includes tests, statistical techniques, and research techniques in cross-sectional and developmental studies. Studies published deal with all levels of education—preschool, elementary, secondary, college, and adult education—and with all age groups.

Manuscripts and correspondence on editorial matters should be addressed to the Editor: Raymond G. Kuhlen, Department of Psychology, Syracuse University, 123 College Place, Syracuse 10, New York.

All manuscripts should be submitted *in duplicate* to facilitate referral to advisory editors and must be prepared to conform to the style and procedures described in the *Publication Manual of the American Psychological Association*. Because of the great demand for journal space, authors should give special attention to reducing the length of manuscripts while maintaining clarity of expression. Ordinarily, papers should not exceed 15 typed pages. Accepted manuscripts are published in order of their receipt. An author may, however, request early publication of accepted manuscripts if he is willing to assume the cost of printing the necessary extra pages.

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Promotion Manager
REBECCA S. HAHN, Editorial Assistant

Communications including subscriptions, orders of back issues, and changes of address should be addressed to the American Psychological Association, 1333 Sixteenth Street, N.W., Washington 6, D. C. Address changes must reach the APA by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Annual subscription: $8.00 (Foreign $8.50). Single copies, $1.50.

PUBLISHED BIMONTHLY BY:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Mt. Royal & Guilford Aves., Baltimore 2, Md.
and 1333 Sixteenth Street N.W., Washington 6, D. C.

Second-class postage paid at Baltimore, Md.
Copyright, 1959 by the American Psychological Association, Inc.

THE JOURNAL OF

EDUCATIONAL PSYCHOLOGY

Volume 50 June, 1959 Number 3

GOAL-SETTING BEHAVIOR AS A FUNCTION OF SELF-ACCEPTANCE, OVER- AND UNDERACHIEVEMENT, AND RELATED PERSONALITY VARIABLES

JAMES V. MITCHELL, JR.

University of Texas

It has never been doubted that goal-setting behavior is strongly influenced by the organization of personality, and several investigators have attempted to ferret out the significant variables of personality that are the major determinants of goal-setting behavior (Cohen, 1954; Frank, 1935; Gardner, 1940; Gould & Kaplan, 1940; Gruen, 1945; Sears, 1941). This has not always been an easy task, and it becomes clear to most that the actions and interactions of variables influencing goal-setting are quite complex and not always subject to simple explanations. In the present study it was not the author's purpose to identify all of the personality variables that influence goal-setting behavior, but rather to analyze the patterns of goal-setting behavior and related personality characteristics of groups that were systematically varied with respect to the two variables of self-acceptance and under- and overachievement.

Procedure

Subjects

Ss were 100 female college students who were majoring in elementary and secondary education and who were either freshmen or sophomores at the time of the investigation.

Measures

The goal-setting behavior of these *Ss* was studied in the actual classroom situation. *Ss* were all members of the author's televised course in elementary educational psychology and hence received exactly the same instruction, except for short discussion periods led by graduate assistants (Miami University, 1957; Mitchell, 1958). Objective examinations were administered at the end of the fifth, eleventh, and sixteenth weeks, and immediately before each examination the subjects were asked the following questions in the order indicated: (*a*) What grade can you reasonably hope to attain on this examination? (*b*) What grade do you actually expect to get?

At the very beginning of the course the *Ss* had been asked the same questions for the entire course, and this was also available for analysis. Also recorded for each *S* was the actual grade received on each examination and for the course as a whole. *Ss* were allowed to use "+" and "−" in setting their goals, and values from zero through nine were assigned to the letter grades A and higher, B+, B, B−, C+, C, C−, D+, D & D−, and F respectively, the lower values being assigned to the higher grades. With these numerical conversions one could represent the difference between an *S*'s level of aspiration and actual grade obtained by subtracting the corresponding value of her actual grade from the value for her level of aspiration, e.g., LA = B or 2; actual grade = C− or 4; difference = −2. Other operations similar to this were also performed. In

computing grades for the course the author made certain that the proportions of *Ss* receiving grades in each of the ten categories described above were the same for every examination, thus keeping this factor constant throughout the entire course. Both grades and aspiration estimates had reasonably normal distributions for each examination, although the means of the aspiration estimates, of course, were consistently higher than the comparable means of the actual grades obtained.

"Self-Acceptance" was measured by means of Bills' Index of Adjustment and Values (Bills, 1951; Bills, 1953). The original Index consists of 49 adjectives, and for each of these the *S* is to rate himself on a five point scale with respect to (*a*) how much of the time the trait is characteristic of him (e.g., "optimistic"), (*b*) how the *S* "feels about" himself as described in the first operation (i.e., I ... [1]—very much dislike ... [2]—dislike ... [3]—neither dislike nor like ... [4]—like ... [5]—like very much ... being as I am in this respect), (*c*) how much of the time the *S* would like to have more of the trait in question (again on a five-point continuum). To this original Index were added three new adjectives related to desire to achieve: "industrious," "hard-working," and "motivated."

In computing an index of under- and overachievement the ACE scores for all *Ss* were normalized and then converted to standard score form. The numerical grade average for each *S* for the semester in which the study was conducted was also computed, and these were also normalized and converted into standard score form. The index of under- and overachievement for each *S* was then defined as the *S*'s *Z* score for her semester grade average minus her *Z* score for her ACE performance.

To analyze the related personality characteristics of the variables described above, a personality inventory was administered that included all of the items of the Taylor Manifest Anxiety Scale (Holtzman, 1952; Taylor, 1953), some additional MMPI items concerned with general metal health, and some items of the author's own devising that were related to the student's desire to achieve.

Design

A preliminary analysis of data in which the Bills' "self-acceptance" scores were plotted against the indices of under- and overachievement yielded a nonsignificant correlation close to zero but revealed a "butterfly" configuration with significantly greater variability in achievement-intelligence ratios for the "self-acceptant" and "self-rejectant" extremes of the Bills' scale than for the moderately acceptant middle values. On the basis of these findings it was hypothesized that the underachievers and overachievers at either end of the Bills' scale were psychologically distinct groups that would exhibit different kinds of goal-setting behavior and different kinds of related personality characteristics.

A "self-acceptant" group ($N = 60$) and a "self-rejectant" group ($N = 40$) were then selected from either extreme of the Bills' distribution on the basis of observable differences in variability between them and the more constricted middle group, and only these extreme groups were used in the major analysis. *Ss* in the "self-acceptant" group had total scores of 170 and above on the Bills' "self-acceptance" scale, while the "self-rejectant" group had total scores of 149 and below. Both of these groups were further subdivided into "overachievers" and "underachievers," with those having positive index values assigned to the former group and those with negative index values assigned to the latter. There were thus four separate groups subject to the analysis: (*a*) the self-acceptant underachievers ($N = 31$), (*b*) the self-acceptant overachievers ($N = 29$), (*c*) the self-rejectant underachievers ($N = 16$), and (*d*) the self-rejectant overachievers ($N = 24$). The median ACE

percentile ranks for each group, based upon the national norms, were 70, 55, 79, and 35, respectively. A two-way analysis of variance with unweighted means was applied to the normalized Z scores of the ACE and yielded an F of .86 (N. S.) between the self-acceptant and self-rejectant groups and an F of 21.95 between the underachievers and overachievers, an expected difference which was significant ($p < .001$). The numerical grade point averages for the semester in which the study was conducted, based upon a conversion system for which A = 4, B = 3, C = 2, etc., were 2.14, 2.92, 2.34, and 2.73, respectively. When the same analysis of variance technique was applied to these grade data, the resultant F of .01 between the self-acceptant and self-rejectant groups was not significant, while the F between the underachievers and overachievers was 26.50, another expected difference that was significant ($p < .001$). Interactions for both the ACE and the grade point averages were insignificant.

For each of the four subgroups of the study the following goal-setting data were secured:

1. Mean level of aspiration (LA) for each examination (from responses to "What grade can you reasonably hope to attain on this examination?").
2. Percentages of students actually achieving their levels of aspiration for each examination.
3. Mean grade expectancy (Gr. Exp.) for each examination (from responses to "What grade do you actually expect to get on this examination?").
4. Mean of the differences between each *S*'s level of aspiration and her expected grade (i.e., LA − Gr. Exp.) for each examination.
5. Mean of the differences between each *S*'s level of aspiration and her actual grade obtained (i.e., LA − Gr.) for each examination.
6. Mean of the differences between each *S*'s expected grade and her actual grade obtained (i.e., Gr. Exp. − Gr.) for each examination.
7. Mean of the differences between each *S*'s actual grade obtained on the last examination and her level of aspiration for the next examination (i.e., $Gr_1 - LA_2$ and $Gr_2 - LA_3$).
8. Mean level of aspiration for the course as a whole, taken at the very beginning of the course (LA_c).
9. Mean of the differences between each *S*'s level of aspiration for the course and the final grade received for the course (i.e., LA_c − Fin. Gr.).
10. Percentages of students actually achieving their levels of aspiration for the course as a whole.

All four subgroups were also compared with respect to their performance on the personality inventory and the self-descriptive section of the Bills' scale. In the former case the percentages of *Ss* responding "Yes" to each item were computed for each group. Items showing large intergroup differences were subject to tests of significance and further analysis. In the latter case, choices for the self-descriptive operation were dichotomized into two groups, with choices 1-*Seldom* is this like me, 2-*Occasionally* this is like me, and 3-*About half of the time* this is like me being placed in the one group; and choices 4-*A good deal of the time* this is like me, and 5-*Most of the time* this is like me being placed in the other group. Percentages of *Ss* in each of these two categories were computed for the four separate subgroups. Items showing large intergroup differences were subject to tests of significance and further analysis.

Results

Table 1 reveals the goal-setting characteristics of the four subgroups. The numbered subscripts 1, 2, and 3 refer to the first, second, and third examinations of the course, while the letter subscript "c" refers to aspiration estimates or grades for the course as a whole. A two-way analysis of variance was employed for all comparisons, using the method of unweighted means as described by Snedecor (1956, p. 385). Interaction effects were omitted from the table, since the only significant

interaction was that for LA_1, which had an F of 4.30 ($p < .05$). Since some previous studies indicated the possibility of significant differences in LA variances between the self-acceptant and self-rejectant groups and between the underachieving and overachieving groups, the F test was applied for both of these comparisons for LA_1, LA_2, LA_3, $LA_1 - Gr._1$, $LA_2 - Gr._2$, $LA_3 - Gr._3$, $Gr._1 - LA_3$, and $Gr._3 - LA_3$. With but one exception no significant differences in variances appeared for any of these comparisons. The one exception involved LA_1, where it appeared that overachievers were significantly ($p < .05$) more variable in setting their first level of aspiration than underachievers.

Inspection of Table 1 discloses some interesting differences in goal-setting behavior for the four subgroups. Within the LA — Gr. category it can be observed that the overachievers come significantly closer to their level of aspiration than the underachievers for Exams 1 and 3, for the Total LA — Gr. averages, and for the course as a whole, and that on Exam 3 the self-rejectant group comes significantly closer to their level of aspiration than the self-acceptant group. Although the latter trend is observable for Exams 1 and 2 as well, it is not statistically significant for these exams, a circumstance which can probably be related to the considerable discrepancy in LA — Gr. for the self-rejectant underachiever in the early part of the semester and the subsequent progressive decrease in LA — Gr. discrepancy toward the end of the semester. This progressive decrease is, in turn, probably attributable to the fact that the self-rejectant underachiever readjusts her LA downward relatively more often and more severely than the others during the semester, as can be observed from the LA data. This is in contrast to the expected grades, which remain fairly constant throughout the semester for all groups.

Differences between the goal-setting behaviors of the sef-acceptant and self-rejectant groups are even more clearly revealed for the Gr. Exp. — Gr. data. Here it is evident that the actual exam grades obtained are consistently lower than the expected grades for the self-acceptant groups, while the self-rejectant groups either have less of a negative discrepancy or actually obtain grades higher than expected, a difference which is significant at the .05 level for Exams 2 and 3 and for the Total Gr. Exp. — Gr. averages. It is also interesting to note that the order of magnitude of the Gr. Exp. — Gr. discrepancies is always the same: self-acceptant underachiever (greatest), self-acceptant overachiever, self-rejectant underachiever, and self-rejectant overachiever (least).

Some other interesting differences are revealed by the Gr. — LA data. The discrepancy between present level of aspiration and previous grade obtained is significantly higher for the underachievers for $Gr._1 - LA_2$. When next computed, however, the picture is somewhat different, for with $Gr._3 - LA_3$ the self-rejectant group now has a significantly smaller discrepancy between present LA and previous grade than the self-acceptant group. Both trends are present for both sets of data, but each attains statistical significance only once. The reason for this can be traced to the behavior of the self-rejectant underachievers and the self-acceptant overachievers. The self-rejectant underachievers are the only ones for whom the $Gr._3 - LA_3$ discrepancy is *less* than the $Gr._1 - LA_3$ discrepancy, despite the fact that they shared with others an appreciable increase in percentage of LAs obtained from LA_1 to LA_2. Quite in contrast is the considerable increase from $Gr._1 - LA_2$ to $Gr._2 - LA_3$ for the self-acceptant overachievers. These two changes are sufficient to change the picture of statistical significance.

Inspection of Table 1 also discloses the following additional information:

1. There is a general trend for LA_2 to be lower than LA_1, probably as a result of the

TABLE 1

COMPARISONS OF THE GOAL-SETTING BEHAVIOR OF SELF-ACCEPTANT UNDERACHIEVERS, SELF-ACCEPTANT OVERACHIEVERS, SELF-REJECTANT UNDERACHIEVERS, AND SELF-REJECTANT OVERACHIEVERS

Operation[a]	Self-Acceptant		Self-Rejectant		*F* for Self-Acceptance and Rejection	*F* for Under- and Over-Achievement
	Under-Achievers	Over-Achievers	Under-Achievers	Over-Achievers		
LA_c	1.17	.71	1.13	1.35	1.99	.32
LA_1	2.73	1.54	2.27	2.43	.44	2.51
LA_2	3.13	1.82	3.00	2.70	.91	4.21*
LA_3	2.57	1.39	3.07	2.22	1.79	4.18*
LA_1 − Gr. Exp.$_1$	−1.13	−.75	−2.13	−1.04	6.40*	8.31**
LA_2 − Gr. Exp.$_2$	−1.03	−1.14	−1.87	−1.17	3.25	1.50
LA_3 − Gr. Exp.$_3$	−1.10	−.89	−1.33	−.91	.31	1.99
LA_1 − Gr.$_1$	−2.33	−1.36	−2.47	−1.04	.04	7.32**
LA_2 − Gr.$_2$	−1.60	−1.32	−1.60	−.39	1.02	2.62
LA_3 − Gr.$_3$	−2.47	−1.21	−1.33	−.70	5.02*	6.59*
Total LA − Gr. (Exams 1 + 2 + 3)	−6.40	−3.89	−5.40	−2.13	2.35	11.86**
LA_c − Fin. Gr.	−3.77	−1.79	−3.33	−1.57	.80	25.56**
% Achieving LA_1	13	43	7	35		
% Achieving LA_2	33	36	47	57		
% Achieving LA_3	10	39	20	61		
% Achieving LA_c	0	25	7	30		
Gr. Exp.$_1$ − Gr.$_1$	−1.20	−.61	−.33	.00	2.81	1.08
Gr. Exp.$_2$ − Gr.$_2$	−.57	−.18	.27	.78	4.30*	1.08
Gr. Exp.$_3$ − Gr.$_3$	−1.37	−.32	.00	.22	5.91*	2.61
Total Gr. Exp. − Gr. (Exams 1 + 2 + 3)	−3.14	−1.11	−.06	1.00	8.71**	2.90
Gr.$_1$ − LA_2	1.93	1.07	1.73	.78	.32	4.31*
Gr.$_2$ − LA_3	2.17	1.75	1.53	.87	4.12*	2.08
Final Grade	4.93	2.50	4.46	2.91	.00	19.81**

Note.—Numerical subscripts 1, 2, and 3 refer to the 1st, 2nd, and 3rd examinations of the course; latter subscript "c" refers to the course as a whole.

[a] Key to abbreviations used:
LA: Level of Aspiration
Gr. Exp.: Grade Expectancy
Gr.: Grade actually obtained
Fin. Gr.: Final Grade in course

* Significant at .05 level.

** Significant at .01 level.

niggardly proportion of *Ss* attaining LA_1. This is followed by an increase in the percentages of *Ss* attaining LA_2, which is followed in turn by an increase in LA_3. (An exception already noted is that for the self-rejectant underachiever.)

2. The level of aspiration for the course is always higher than LA_1, even though no

formal evaluation has occurred in the meantime. Aparently just a few weeks of course work encourages more realistic goal-setting.

3. The self-rejectant overachievers always have the least discrepancy between level of aspiration and grade actually obtained.

4. The self-rejectant overachievers are consistently most likely to receive an exam grade that fulfils or surpasses their stated expectations.

5. The self-rejectant overachievers always have the least discrepancy between previous grade and present level of aspiration (which partially explains 3 and 4 above).

6. The self-acceptant underachievers always have the greatest discrepancy between previous grade and present level of aspiration.

7. The self-acceptant underachievers always have the greatest discrepancy between expected grade and grade actually obtained.

8. For the first exam the discrepancy between level of aspiration and grade expectancy is significantly greater for the underachievers as opposed to the overachievers and significantly greater for the self-rejectant as opposed to the self-acceptant. Though these trends are observable again for the second and third exams, they are not significant for these, primarily because of the progressive decrease in LA − Gr. Exp. discrepancy for the self-rejectant underachievers. The latter, however, are observed to have the greatest discrepancy for all three exams, despite this progressive decrease.

Tables 2, 3, and 4 show some of the personality characteristics associated with membership in each of the four separate subgroups. Table 2 indicates that scores on the Taylor Manifest Anxiety Scale were significantly higher for the self-rejectant *Ss* as opposed to the self-acceptant *Ss* and for the overachievers as opposed to the underachievers. The cumulative effect of these two trends is to reveal the self-rejectant overachiever as the highest-ranking on the anxiety scale and the self-acceptant underachiever as the lowest. It is interesting to note that the order of magnitude for these scores is exactly the same as that for the Gr. Exp. − Gr. data for all three exams, and also for the $Gr_{-2} - LA_2$ data as well.

TABLE 2

SCORES ON THE TAYLOR MANIFEST ANXIETY SCALE FOR THE FOUR SUBGROUPS

Self-Acceptant		Self-Rejectant		*F* Test for Self-Acceptance and Rejection		*F* Test for Under- and Over-Achievement	
Under-Achievers	Over-Achievers	Under-Achievers	Over-Achievers	*F*	*p*	*F*	*p*
11.55	14.72	19.88	23.29	31.18	.001	4.72	.05

Note.—The interaction effect was insignificant and therefore omitted.

Table 3 shows those items in the personality inventory that produced statistically significant discriminations among the four subgroups of the study. For each of these items the numbers of *Ss* answering "Yes" and "No" were computed for each subgroup, and a chi square value was computed for the resulting 2 by 4 table. Table 4 shows those items in Bills' Index of Adjustment and Values that produced statistically significant discriminations. These same operations had been carried out for the Bills' scale, the responses being dichotomized as described in a previous section. However, certain problems arose with the Bills' scale that necessitated a somewhat different type of analysis. It was evident that statistically significant differences would appear for items on the Bills' scale by virtue of the fact that the self-acceptant *Ss* would naturally rate themselves higher on positive traits than would the self-rejectant *Ss*. The same trend is also observable for the data of Table 3, but here the effect does not seem to be as great. But for the Bills' scale it was decided to *rank* the items within each group according to the percentage of positive responses (Alternatives 4 & 5) and then to add these rank order data to the table for interpretive purposes. The ranked data constitute a self-representation by each group in terms of characteristics seen as more typical or less typical of

TABLE 3

INVENTORY ITEMS HAVING SIGNIFICANT CHI SQUARE VALUES FOR THE FOUR SUBGROUPS

Item	Percentage Answering "Yes"				X^2
	Self-Acceptant		Self-Rejectant		
	Under-Achievers	Over-Achievers	Under-Achievers	Over-Achievers	
When I'm faced with difficult task I generally "give up" too easily.	10	7	50	17	15.37**
I wish I were more successful as a student.	90	66	94	92	10.27*
I try to do well in anything I undertake.	97	90	56	92	16.36**
I am more sensitive than most other people.	20	46	38	58	8.72*
I frequently find myself worrying about something.	47	69	50	79	7.53a
I find the standards I set for my own work are usually too high for me to attain.	40	14	19	50	10.26*
I wish I could be as happy as others seem to be.	16	31	31	58	10.61*
I am usually calm and not easily upset.	97	79	56	46	19.61**
I usually expect to succeed in things I do.	87	100	88	75	7.78a
I have periods of such great restlessness that I cannot sit long in a chair.	30	21	56	54	9.48*
I have sometimes felt that difficulties were piling up so high that I could not overcome them.	50	55	69	79	8.39*
I find it hard to keep my mind on a task or job.	16	14	63	38	17.09**
I am a high-strung person.	0	21	31	38	13.26**
Life is a strain for me much of the time.	0	7	19	33	14.74**
I have strong motivation to succeed in my work.	70	79	44	92	12.09**
At times I think I am no good at all.	30	34	81	71	18.00**
I am certainly lacking in self-confidence.	20	24	50	75	21.05**
I am entirely self-confident.	40	24	19	0	12.52**
I brood a good deal.	7	10	6	29	8.34*
I'll never be able to live up to my parents' expectations for me.	7	10	38	29	9.61*
I am liked by most people who know me.	80	97	81	88	7.13a

a Approaches statistical significance where $p.05 = 7.82$; included here for its interpretive value in discriminating the subgroups.

* Significant at .05 level.

** Significant at .01 level.

them as individual members of the group. Three of the traits—"optimistic," "stubborn," and "competitive"—were included in Table 4 despite the fact that they were not statistically significant, because it was anticipated that these three would exhibit appreciable intergroup differences in *rank* even though statistical significance was absent. This proved to be the case for two of these three traits.

Inspection of Table 3 along with the ranked data of Table 4 discloses some interesting differences between the four subgroups. The self-acceptant underachiever rates herself higher than the others on such traits as happiness, self-confidence, alertness, being less sensitive than others, worrying less, being calm and not high-strung or easily upset, and not finding life much of a strain. But she rates herself low on ambition. The self-acceptant overachiever rates herself high on maturity, worthiness, competence, and industry, considers herself well-liked, and

TABLE 4

ITEMS IN THE "INDEX OF ADJUSTMENT AND VALUES" HAVING SIGNIFICANT CHI SQUARE VALUES FOR THE FOUR SUBGROUPS

Item	Percentage Choosing Alternatives "4" and "5"				χ^2	Item Rank Within Group			
	Self-Acceptant		Self-Rejectant			Self-Acceptant		Self-Rejectant	
	Under-Achievers	Over-Achievers	Under-Achievers	Over-Achievers		Under-Achievers	Over-Achievers	Under-Achievers	Over-Achievers
acceptable	100	90	87	46	29.55**	1	5	1	9
alert	90	72	25	37	27.30**	5	14	15.5	14
ambitious	56	79	31	58	11.58**	20	13	11.5	4
calm	81	69	62	37	13.28**	9	15	3.5	14
competent	81	86	31	37	19.34**	9	7.5	11.5	14
confident	81	66	19	12	34.15**	9	16.5	19.5	23
helpful	81	93	50	54	15.48**	9	3	5.5	6.5
logical	77	66	25	42	14.22**	12.5	16.5	15.5	11
industrious	58	86	25	33	19.47**	18.5	7.5	15.5	17
merry	90	86	62	58	11.12*	5	7.5	3.5	4
mature	90	97	37	37	36.62**	5	1	8.5	14
nervous	10	14	19	37	7.60[a]	23.5	24	19.5	14
optimistic	61	62	50	29	6.00	16.5	19.5	5.5	19
poised	71	62	37	29	11.99**	14.5	19.5	8.5	19
responsible	42	59	12	8	19.05**	21.5	22	22.5	24
hardworking	61	83	19	46	18.86**	16.5	10.5	19.5	9
stable	93	83	44	67	16.23**	2.5	10.5	7	1
studious	42	62	6	46	13.21**	21.5	19.5	24	9
successful	71	81	31	21	21.17**	14.5	12	11.5	21.5
stubborn	10	17	19	21	1.64	23.5	23	19.5	21.5
worthy	77	93	31	29	32.74**	12.5	3	11.5	19
broad-minded	93	93	69	62	13.31**	2.5	3	2	2
competitive	58	62	25	58	6.56	18.5	19.5	15.5	4
motivated	81	86	12	54	29.69**	9	7.5	22.5	6.5

[a] Approaches statistical significance where $p_{.05} = 7.82$; included here for its interpretive value in discriminating the subgroups.

* Significant at .05 level.

** Significant at .01 level.

rates herself extremely high on her expectancy "to succeed in things I do." The self-rejectant underachiever rates herself low on motivation and high on "giving up" too easily. She is more likely to report that she is not studious or "hardworking." She is also more likely to report that she finds it hard to keep her mind on a task and that at times she thinks she is no good at all. In spite of this she ranks calmness and optimism high in her self-ratings. The self-rejectant overachiever rates herself high on ambition, motivation in schoolwork, studiousness, competitiveness, nervousness, sensitivity, and tendency to brood or worry. She rates herself low on confidence, competence, success, acceptability, maturity, and worthiness. She is also more likely to report unhappiness, being high-strung and easily upset, feeling overwhelmed with difficulties, setting standards too high to attain, not expecting to succeed in things she does, and finding life a strain much of the time.[1]

After the data of Table 4 had been ranked to facilitate interpretation, it seemed fruitful to take one further step and compute intergroup rank-difference correlations to determine the degree of similarity and difference between the rankings of the various groups. These data are shown in Table 5. One notable difference here is that the self-rejectant overachiever is *least* similar to all the others for these selected traits. The traits which serve most to distinguish her from the others appear to be the following: strong competitiveness, ambition, studiousness, feelings of unworthiness, acknowledged immaturity, extreme lack of self-confidence, lack of success feelings, and nervousness and anxiety. These data should be considered in the light of the fact that the self-rejectant overachievers are definitely a below average group intellectually, having a median ACE percentile rank of only 35.

[1] It is curious that the self-rejectant overachiever regards *stability* as the trait *most* characteristic of herself, and one suspects that the term is so emotionally loaded that it activates the strongest defensive reactions in those who are the most *un*stable. It is also sufficiently nondescriptive to encourage unthinking (and generally positive) responses.

TABLE 5

INTERGROUP CORRELATIONS FOR THE RANKED DATA OF TABLE 4

		Self-Acceptant		Self-Rejectant	
		Under-Achievers	Over-Achievers	Under-Achievers	Over-Achievers
Self-Acceptant	Under-Achievers		.71	.65	.40
	Over-Achievers			.52	.39
Self-Rejectant	Under-Achievers				.33
	Over-Achievers				

DISCUSSION

Data have been presented concerning the goal-setting patterns of the four subgroups and also the personality characteristics distinguishing the subgroups. An attempt will now be made to interpret these empirical relationships and portray their meaning for each of the groups involved.

Some significant trends are evident for the data as a whole. By inspection it is evident that self-rejection and manifest anxiety are related. (An earlier investigation revealed a significant correlation of .41 between these two variables.) Both of these are related in turn to a cautious and conservative pattern of goal-setting, the most anxious, self-rejectant *Ss* setting their new LAs closest to the grades obtained on the previous exam. It is probably because of this that these *Ss* more often exceed their stated grade expectations than is typical of the less anxious, self-acceptant subjects, whose actual

grades are characteristically lower than their stated grade expectations. One explanation for this is that the self-rejectant person tends to be relatively unconfident and pessimistic when predicting her grade, and she is so conservative in her predictions that she often performs better than she expects. The self-acceptant person, on the other hand, may expect more and generally get less. An alternative but related explanation would be that the already anxious, self-rejectant person is consciously or unconsciously aware that critical nonattainment of goals would create additional tension and anxiety with which she could not cope, and to prevent this from occurring she lowers her aspirations accordingly. It is probable that both of these operate in some measure.

Operating within this general trend are several specific tendencies that can be noted for each of the separate subgroups:

The Self-Acceptant Underachiever: There is a strong suggestion in the data that the self-acceptant underachiever either fulfils her ego needs in areas other than the academic or that she can somehow achieve self-satisfaction and contentment more easily (and with less justification) than the average person. Her calm, unworried, self-confident kind of temperament permits her to aspire to grades that are further beyond actual previous attainment than those for any other group. However, she is not especially ambitious or willing to exert much effort in order to achieve her goals, and she therefore also falls further short of her expectations than those in any other group. She often recognizes that she has not attained the standards she set for herself, but nonattainment is apparently of little concern to her, since she is the least anxious of all those represented in the study. Doubtlessly she is well-adjusted, but she falls far short of realizing her potentialities as an individual.

The Self-Acceptant Overachiever: The self-acceptant overachiever can probably best be characterized as the well-adjusted, "good" student. Although often of mediocre intelligence, she consistently aspires to the highest grades, and her actual grades are generally the highest obtained by any of the subgroups. She is a hard working, industrious person whose diligence has assured success in the past and who therefore just expects further success in the future. She probably obtains a large measure of ego-satisfaction through her academic work. She can put pressure on herself when she needs it most; toward the end of the course she considerably increased the gap between the last exam grade and her current LA and then went on to come as close or closer to this new aspiration than she ever had before. She sees herself as a mature, worthy, well-liked person who is competent to deal with life's challenges and can do so without undue nervousness or anxiety. Unlike the self-rejectant overachiever, her motivational system is not characterized by a strong sense of competitiveness.

The Self-Rejectant Underachiever: The self-rejectant underachiever is, without doubt, the least motivated of all. She has the highest median ACE score of all four subgroups; yet she is more likely to admit that she does not try too hard, that she tends to "give up" too easily, and that she cannot keep her mind on her work. She ranks herself relatively high in "optimism," however, and both this optimism and her acknowledged tendency to give up too easily are reflected in the manner in which she sets her goals. She starts out bravely enough and demonstrates her initial "optimism" by setting an LA that is further from her expected grade than those for other groups, and this LA − Gr. Exp. gap continues to be greatest on the succeeding exams. But only 7% of the self-rejectant underachievers achieve LA_1, and there ensues a rather precipitous drop in LA for succeeding exams. The drop is demonstrably greater from LA_1 to LA_2 than for any other group, and from LA_2 to LA_3 all the other groups are actually increasing their LAs appreciably, while that for the self-rejectant underachiever continues to decrease. This tendency to pull down the LA, obviously related to present lack of motivation as well as initial overestimation, is also indicated in the data relating current LA with previous exam grade, where it is observed that the discrepancy *increases* over time for all the other groups but actually *decreases* for the self-rejectant underachiever. This has the effect of progressively decreasing the LA − Gr. Exp. and LA − Gr. discrepancies, but only because of the drop in LA and certainly not because of increased motivation and willingness to work hard. As a matter of fact, the evidence could most reasonably be interpreted as pointing to a progressive decrease in motivation throughout the semester. Perhaps it is because of

her unrealistic goal-setting behavior, her lack of motivation, and her chronic nonattainment that the self-rejectant underachiever feels at times that she is "no good at all." This may also help to explain her high manifest anxiety score. Why, then, does she also characterize herself as relatively calm, poised, and not inclined to brood or worry? One suspects that the self-rejectant underachiever expends much psychic energy in defense maneuvers of one kind or another; it is very possible that her great LA − Gr. Exp. discrepancies connote level of aspiration behavior that is essentially defensive in nature—more so than for any other group. It would not be surprising if a psychic life characterized by strong repressive tendencies and defensive reactions were to result in manifest anxiety and equally manifest underachievement.

The Self-Rejectant Overachiever: Probably the most depressing picture of all is presented by the self-rejectant overachiever. Her self-description indicates that she is more unhappy, more nervous, more high-strung, and more easily upset that those in the other subgroups; and her manifest anxiety score, which is the highest of all four subgroups, confirms this self-description. She feels incompetent, unsuccessful, unworthy, and completely lacking in self-confidence; and she is often inclined to brood over her inadequacies and her troubles and to be utterly overcome by life in general. Her anxiety and lack of self-confidence are reflected in her goal-setting behavior, for she consistently sets her goals more cautiously and conservatively than those in the other subgroups. She sets her LAs far closer to previous grade obtained than the others, and she always has the least discrepancy between LA and grade obtained. She also more often exceeds her expected grade than anyone else. She acknowledges that she is ambitious, studious, highly motivated, and strongly competitive; and the combination of this with her cautious goal-setting behavior is probably responsible for the smaller LA − Gr. discrepancy and the greater tendency to exceed grade expectations. Characteristically, however, she reports more often than the others that "I find the standards I set for my own work are usually too high for me to attain" *despite* the fact that she comes closer to attaining them than anyone else! Faced with her deep-seated feelings of unworthiness, the self-rejectant overachiever apparently spends most of her time in a struggle to achieve some shred of self-respect. This struggle takes the form of a desperate and highly motivated effort to achieve academic excellence. This must be a very difficult task for many members of this group, since the median ACE percentile rank is only 35, and the combination of mediocre ability and desperate desire for academic achievement must only create further tension and anxiety that produce additional misery in an already unhappy life. Doubtlessly it is because of this extreme emotional investment in goal-attainment that the self-rejectant overachiever sets her goals so conservatively and then tries so desperately to attain them. Her sense of personal unworthiness is deeply felt, and that is why she must put forth so much effort to convince herself and others that she is a worthy person after all.

Summary

The purpose of this investigation was to analyze the patterns of goal-setting behavior and related personality characteristics of four groups: self-acceptant underachievers, self-acceptant overachievers, self-rejectant underachievers, and self-rejectant overachievers. Self-acceptance and rejection were measured by the Bills' Index of Adjustment and Values. *Ss* were 100 female college students enrolled in a course in educational psychology. Three examinations were given during the course, and before each of these each *S* was asked to indicate the grade she hoped to attain on the exam and the grade she actually expected to get. In addition to the Bills' scale a personality inventory was administered that included the items of the Taylor Manifest Anxiety Scale and other items relating to mental health and desire to achieve. Among other results reported, it was noted that self-rejectant *Ss* had a significantly smaller discrepancy between previous grade and present level of aspiration, and there was also an accompanying trend for the self-rejectant *Ss* to exceed their expected grades while the self-acceptant consistently overestimated theirs. Since underachievement was also associated with overestimation of grade, the self-acceptant underachiever led all four groups in gross overestimation, while

the self-rejectant overachiever led the others in either achieving or exceeding the expected grade. Significant differences in manifest anxiety were revealed, with the underachievers and self-rejectant *Ss* exhibiting greater anxiety, a condition that was reflected in the more cautious and conservative goal-setting that characterized these groups. Careful analysis of the items in the Bills' scale and in the personality inventory revealed many interesting differences between the four subgroups.

REFERENCES

Bills, R. E. An index of adjustment and values. *J. consult. Psychol.*, 1951, **15**, 257–261.

Bills, R. E. A comparison of scores on the index of adjustment and values with behavior in level of aspiration tasks. *J. consult. Psychol.*, 1953, **17**, 206–212.

Cohen, L. D. Level of aspiration behavior and feelings of adequacy and self-acceptance. *J. abnorm. soc. Psychol.*, 1954, **49**, 84–86.

Frank, J. D. Some psychological determinants of the level of aspiration. *Amer. J. Psychol.*, 1935, **47**, 285–293.

Gardner, J. W. The relation of certain personality variables to level of aspiration. *J. Psychol.*, 1940, **9**, 191–206.

Gould, Rosalind, & Kaplan, N. The relationship of level of aspiration to academic and personality factors. *J. soc. Psychol.*, 1940, **11**, 31–40.

Gruen, E. W. Level of aspiration in relation to personality factors in adolescents. *Child Develpm.*, 1945, **16**, 181–188.

Holtzman, W. H., Calvin, A. D., & Bitterman, M. E. New evidence for the validity of Taylor's manifest anxiety scale. *J. abnorm. soc. Psychol.*, 1952, **47**, 853–854.

Miami University. *Second report: Experimental study in instructional procedures.* Oxford, Ohio: Miami Univer., 1957.

Mitchell, J. V., Jr. Teaching educational psychology by TV. *Improving College and University Teaching*, 1958, **6**, 90–95.

Sears, Pauline S. Level of aspiration in relation to some variables of personality: clinical studies. *J. soc. Psychol.*, 1941, **14**, 311–336.

Snedecor, G. W. *Statistical Methods.* Ames, Iowa: Iowa State Coll. Press, 1956.

Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.

Received October 16, 1958.

JOURNAL OF EDUCATIONAL PSYCHOLOGY
Vol. 50, No. 3, 1959

PREDICTING LEADERSHIP RATINGS FROM HIGH SCHOOL ACTIVITIES[1]

JOHN D. KRUMBOLTZ
Michigan State University

RAYMOND E. CHRISTAL AND JOE H. WARD, JR.
Personnel Laboratory, Wright Air Development Center

Can a self-report questionnaire about high school activity participation be used to predict later leadership peer ratings? A recent review of the literature (Krumboltz, 1957b) concerning the relationship of high school activities to leadership criteria revealed no conclusive evidence one way or the other. There was some evidence that college activities were predictive of future leadership success, but the studies dealing with high school activities were of such a nature that no valid conclusions could be drawn. More recently, a study (Krumboltz & Christal, 1957) at the Air Force Academy revealed correlations ranging from .21 to .28 between high school activity participation and leadership ratings. The present study was designed to determine whether high school activity participation can be used to predict future leadership ratings within the aviation cadet population. It is an attempt to verify the findings of the Air Force Academy study using a different population. However, the present study differs in three respects. First, a special inventory form was developed by which the examinee records his high school activity participation on an IBM Mark Sense card. Secondly, data for individuals from large and small high schools were analyzed separately. Finally, individual items were selected and combined by an iterative multiple regression technique to produce the highest possible relationship with the criterion.

The establishment of any positive relationship between high school activity participation and leadership criteria does not, of course, necessarily indicate a cause and effect relationship. If a successful prediction can be made, this fact has important implications in a selection program whether or not leadership ability may have resulted from the training received in high school activities.

Sample

The total sample consisted of 956 aviation cadets in preflight training at Lackland Air Force Base, Texas. Of these, 857 graduated from preflight training, while 99 were eliminated for a variety of reasons. Since students from smaller high schools generally have available a fewer number of high school activities (although perhaps more opportunity to participate in those activities which are available), students from large and small schools were analyzed separately. Each cadet was asked to indicate approximately how many persons there were in his graduating class including midyear and summer session graduation. If he reported 99 or less, he was arbitrarily considered to be from a "small" high school. If he reported 100 or more, he was arbitrarily considered to be from a "large" high school.

Six aviation cadet classes were involved in this study. The first three, Classes V-15,

[1] This report is based on work done under ARDC Project No. 7719, Task No. 17009, in support of the research and development program of the Personnel Laboratory, Wright Air Development Center, Lackland Air Force Base, Texas. Additional support was received from an All-University Research Grant from Michigan State University.

A-16, and B-17, who entered training in June and July of 1956, were grouped together and are hereinafter referred to as Group A. The three classes who entered training in July and August immediately after Group A, Classes C-18, D-19, and E-20, are hereinafter referred to as Group B.

Thus, four subsamples were formed out of the 857 graduates for analysis purposes:

1. Small high school, Group A, $N = 162$;
2. Small high school, Group B, $N = 135$;
3. Large high school, Group A, $N = 306$;
4. Large high school, Group B, $N = 254$.

The High School Activities Inventory

The High School Activities Inventory was devised to measure the extent and nature of each individual's participation in his high school's extracurricular (or cocurricular) activity program. Since extracurricular programs vary considerably in different parts of the country and among different high schools in the same geographical region, it was necessary to devise a questionnaire that could be interpreted equally well by anyone attending high school anywhere in the United States or its territories. Fortunately, it was possible to obtain the records of applicants to the Air Force Academy Class of 1959. Each applicant had written out a list of all his extracurricular activities. These lists were then consolidated into a master list of 226 separate activities. The activities ranged from polo to chess and many of them were listed by only one individual. By combining very specific activities into more general categories and by eliminating rare types of activities, it was possible to reduce the list to a more manageable 44 activities which were included in the final inventory. Actually there were 70 items since the extent of one's participation was often included as a separate item under one given activity. For example, participating in football and winning a letter in football were two separate items.

Each item takes the form of asking how many years an individual participated in a given activity or how many years he received a certain honor. The initial instructions directed the examinees to consider only the last three years of their high school career in answering how many years they participated. Thus, there were only four possible answers to each item: 0, 1, 2, and 3. In this way individuals who attended four-year high schools had no "time" advantage over individuals who attended three-year high schools.

It was considered possible that some cadets might try to exaggerate the extent of their extracurricular participation although there was no reason why they should. To discourage this possibility each cadet was asked to list the names and locations of the high schools he attended, the names of the high school principals, and the names of three teachers at each high school who could verify his activities. Although no attempt was made to follow up this information, it was believed that the possibility of verifying a cadet's responses made it less likely that he would exaggerate them.

Responses and identifying data were recorded on mark sense cards (IBM 850259MS-O). A list of 20 of the 70 items may be found in Table 1 in summarized form. For the most part each item is self-explanatory if it is understood that each activity in Table 1 is preceded by "How many years did you participate in . . ." or "How many years did you receive this honor . . . ."

To determine how well the inventory covered the wide variety of possible activities the first 200 cadets to answer the inventory were asked to write on the back of the inventory any activities in which they had participated but which were not covered in the inventory. A total of 105 comments were received including two or more comments from some cadets. 85 of the comments were judged to be irrelevant for a variety of reasons, while 20 were considered as legitimate criticisms of omis-

sions in the inventory. The irrelevant comments concerned activities the questionnaire was not designed to cover such as intramural athletics, high school fraternities, scholastic awards, and church activities. Other irrelevant comments included very specific honors or activities which could not be classified, activities that had already been covered in the inventory directly or indirectly, and excuses for little or no participation. The main area of activity that was omitted by the inventory concerned school service activities. Four persons listed the visual aid department, five were managers of athletic teams, seven served on various school committees, and four listed other miscellaneous school activities.

Criterion

Leadership peer ratings were collected on the aviation cadet sample twice during their preflight training. Each man in a flight was asked to rank-order each other man in his flight on the basis of his leadership potential. The first time was after the first four weeks of training, and the second time was after the first ten weeks of training. Cadets were rank-ordered on the basis of the pooled rankings within each flight, and the rank orders were then converted to T scores. The stability of these rankings over a six-week interval is indicated by correlations for the six classes ranging from .40 to .82 with a median r of about .68.

A number of studies (Hollander, 1957; Trites & Sells, 1957; United States Department of the Army: 1949, 1952) have demonstrated the reliability and validity of peer ratings as an intermediate criterion of future leadership behavior. The results of these and other studies have been briefly discussed elsewhere (Hollander, 1957; Krumboltz, 1957a). It may be concluded from these studies that leadership peer ratings have a substantial relationship to future leadership ratings collected in an on-the-job situation.

Statistical Methodology

Two techniques were used to analyze the data. The first consisted of summing the weighted responses to each item (i.e., adding the number of years recorded for each item) and computing the zero order correlation of this sum with the criterion. This was done separately for athletic items and for nonathletic items for each of the four samples.

The second technique consisted of selecting and weighting items by an iterative multiple regression technique so as to produce the highest possible relationship with the criterion. The basic procedure that was used to obtain the regression coefficients has been described by Greenberger and Ward (1956). This technique starts with all regression coefficients equal to zero. At each iteration the unique correction is computed for each regression coefficient that will maximally increase the multiple correlation. The alteration is made only on the one particular variable which has the largest correction.

For each of the four samples it was first necessary to compute the column vector of validities. The iterative technique then proceeds to indicate the column vectors of intercorrelations that are successively required for the regression computation. As a result, if there are only k predictor variables used out of the 70 original potential predictors, then it is necessary to compute only the 70 by $(k + 1)$ rectangular matrix of correlation coefficients.

The problem associated with "when to stop iterating" is as yet unsolved; however, it has been observed that an F test for the significance of the increase of predictive efficiency gives an approximate indication of a "good" place to stop. After the "best" set of k variables had been selected for each sample, the iterations were continued on the variables that were retained until a least squares solution was reached for the k predictors. The regression equation computed for each of the

TABLE 1

RELATIONSHIP OF HIGH SCHOOL ACTIVITY ITEMS TO THE LEADERSHIP CRITERION

Item No.	Activity	Correlation with Criterion in				Mean No. of Years per Man (N = 857)
		Small HS		Large HS		
		A (N = 162)	B (N = 135)	A (N = 306)	B (N = 254)	
1	Football	.18*	.18	.14*	.19	.80
3	Basketball	.04	.29*	.13	.25*	.56
7	Track	.18	.12	.07	−.02	.46
13	Tennis	.05	−.08	.11*	.00	.11
15	Swimming	.04	.20*	.02	.06	.11
19	Other sport	−.16*	−.02	.03	−.01	.24
23	Newspaper assistant editor	−.15*	−.04*	−.13	.14	.09
25	Newspaper photographer	−.07	−.07	−.17*	.02	.04
34	Debate team	.14*	.18*	.11	−.03	.10
36	Chorus or glee club	.07	.05	−.01	.18*	.52
41	Science club member	−.08	−.23*	−.08	.04	.31
44	Language club member	.00	.16*	.00	.12*	.22
47	Hobby or interest club member	−.15	−.17*	−.04	−.09*	.49
57	Student government president	−.06	.11	.11	.20*	.06
58	Student government: other officer	−.12*	−.07	.08	.14	.14
59	Student government: member	−.03	.09*	.15	.11	.34
63	President of class	−.02	.06	.16	.23*	.14
64	Vice-president of class	.09	.20	.18*	.14	.12
65	Secretary of class	.03	.12	.15*	−.06	.07
69	Outstanding student award	.13*	−.01	−.04	.00	.13

* Items weighted into multiple regression equations.

four samples was then applied to all samples for cross validation.

RESULTS AND DISCUSSION

The correlations of each of 20 items with the criterion for each of the four samples are reported in Table 1.[2] Asterisks beside a correlation indicate that the item was weighted into the multiple regression equation for that sample. The 20 items appearing in Table 1 were selected from among those judged to be of most interest and yielding the most substantial relationships with the criterion.

Positive correlations were consistently obtained for most of the major sports and for athletic honors. In small high schools being assistant editor on the school newspaper received negative weight, while being a member of the debate team received positive weight. In large high schools being president or vice-president of the class was consistently and positively related to

[2] A four-page table for all 70 items and including the extent of participation by eliminees as well as graduates has been deposited with the American Documentation Institute, Order Document No. 5955 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

TABLE 2

ZERO ORDER CORRELATIONS OF ACTIVITY SCORES WITH THE LEADERSHIP CRITERION

Sample	*N*	Correlations[a]				Athletic Score		Nonathletic Score	
		r_{ac}	r_{nc}	r_{nn}	r_{tc}	Mean	*SD*	Mean	*SD*
1. SA	162	.10	−.05	.24**	.02	6.3	6.3	10.8	7.3
2. SB	135	.29**	.02	.28**	.18	5.7	5.2	9.2	6.1
3. LA	306	.18**	.06	.27**	.12	3.2	3.8	7.6	6.5
4. LB	254	.23**	.18**	.42**	.24**	3.6	4.2	6.7	5.5

[a] Subscripts represent the following variables: a, Athletic Activity Score; n, Non-athletic Activity Score; t, Total Activity Score; and c, the leadership criterion.

** Significant at the .01 level.

high leadership ratings, while such consistent relationships were not found for the small high school samples. Being a member or officer in student government was positively related to the criterion in large high schools but not in small high schools.

Many items that one might reasonably expect to be positively related to leadership ratings showed no relationship whatsoever. For example, all of the following were unrelated to the criterion: being a delegate to Boys' State; being president of a career interest club, a hobby or interest club, a language club, or a science club; and being editor of the school newspaper or yearbook. The reasons for this are a matter for conjecture. Perhaps individuals with relatively low leadership ability go into such activities. Being president or editor in a group of nonleaders would then give no indication of one's leadership status in a larger more representative group. Perhaps leadership is specific to the situation.

By summing each individual's responses to items 1–20 and 70, and Athletic Activity Score was computed for each individual. Similarly the summation of each cadet's responses to items 21–69 provided his Nonathletic Activity Score. The sum of these two scores was the Total Activity Score. Table 2 reports the zero order correlations of each score with the criterion and the correlations of the Athletic and Nonathletic Activity Scores for each sample. The correlations of the Athletic Activity Score with the criterion ranged from .10 to .29 and were consistently higher than the correlations of the Nonathletic Activity Score with the criterion which ranged from −.05 to .18.

The multiple regression equations derived from each sample yielded spuriously high multiple *R*s ranging from .36 to .62. When the regression equations derived from each sample were cross-validated on each other sample, the results shown in Table 3 were obtained. It is of special interest to note the cross-validities of the small high school samples on each other and the cross-validities of the large high school samples on each other. Multiple *R*s shrank from .47 to .17 and from .62 to .19 in the small high school samples. In the large high school samples the multiple *R*s shrank from .36 to .19 and from .42 to .14.

TABLE 3

VALIDITIES OF REGRESSION EQUATIONS IN EACH SAMPLE

Weights Derived from Sample	*N*	Validated on Sample				Criterion	
		1	2	3	4	Mean	*SD*
1. SA	162		*.19*	.11	.03	100.7	16.1
2. SB	135	*.17*		.19	.12	98.7	15.8
3. LA	306	.16	.21		*.14*	100.1	14.5
4. LB	254	.11	.29	*.19*		100.1	15.5

Summary and Conclusions

The purpose of this study was to determine whether a self-report questionnaire about high school activity participation can be used to predict later leadership peer ratings. An inventory on high school activity participation was administered to a total of 956 aviation cadets undergoing preflight training at Lackland Air Force Base. The criterion consisted of the summation of *T*-scored leadership peer ratings collected after the fourth and after the tenth week of training. The data were analyzed by computing zero order and multiple correlations between the predictor items and the criterion.

The analysis revealed consistently low but positive correlations between inventory scores and the criterion. The multiple regression equations derived from each subsample yielded multiple *R*s which shrank upon cross-validation on comparable samples to values which ranged from .14 to .19 with a median of .18. Three of the four zero order correlations between the Athletic Activity Index and the criterion were significantly different from zero at the .01 level. Only one of the four correlations between the Nonathletic Activity Index and the criterion reached significance.

On the basis of these results the following conclusions were reached:

1. A self-report inventory on extracurricular participation succeeded in predicting leadership ratings of aviation cadets better than chance, but considerable error still remained in such predictions. The low positive correlations do not, of course, necessarily mean that training received in high school activities produced increased leadership ability since those choosing to take part in high school activities might very well have been above average in leadership ability originally.

2. Although certain types of activities tended to have differential predictive power for people from large and small high schools, in general activities in large and small high schools were about equally predictive of the criterion.

3. Athletic participation and honors were more predictive of future leadership than nonathletic participation and honors.

4. The iterative multiple regression technique did not produce a more valid composite than the simpler zero order correlation technique in this study.

REFERENCES

Greenberger, M. H., & Ward, J. H., Jr. An iterative technique for multiple correlation analysis. *IBM Applied Science Division Technical Newsletter*, 1956, No. 12, 85–97.

Hollander, E. P. The reliability of peer nominations under various conditions of administration. *J. appl. Psychol.*, 1957, **41**, 85–90.

Krumboltz, J. D. Physical proficiency as a predictor of leadership. *USAF Personnel Train. Res. Cent. Res. Rep.*, 1957, No. 57-60. (ASTIA Document No. 126391.) (a)

Krumboltz, J. D. The relation of extracurricular participation to leadership criteria. *Personal guid. J.*, 1957, **35**, 307–314. (ASTIA Document No. 134202.) (b)

Krumboltz, J. D., & Christal, R. E. Predictive validities for first-year criteria at the Air Force Academy. *USAF Personnel Train. Res. Cent. Res. Rep.*, 1957, No. 57-95. (ASTIA Document No. 134218.)

Trites, D. K., & Sells, S. B. Combat performance: Measurement and prediction. *J. appl. Psychol.*, 1957, **41**, 121–130.

U. S. Department of the Army, Personnel Research Section. Follow-up validation of predictor instruments for West Point Classes of 1944, 1945, and 1946 against 1948 ratings on DA AGO Form 67-1. *USA TAGO Personnel Res. Br. tech. Res. Rep.*, 1949, No. 811.

U. S. Department of the Army, Personnel Research Section. Studies of the performance of officers in combat: I. Relationship of West Point measures to later combat effectiveness. *USA TAGO Personnel Res. Br. tech. Res. Rep.*, 1952, No. 969.

Received July 21, 1958.

JOURNAL OF EDUCATIONAL PSYCHOLOGY
Vol. 50, No. 3, 1959

EFFECTS OF SOCIAL RECOGNITION UPON THE EDUCATIONAL MOTIVATION OF TALENTED YOUTH

DONALD L. THISTLETHWAITE[1]
National Merit Scholarship Corporation

A persistent problem in American education—made more imperative by recent concern over Sputnik—is that of stimulating able students to realize their educational potential. What can be done to motivate talented youth to seek further specialized training? Because careers in science and education do not usually pay well, talented students are frequently diverted into careers which, though more lucrative, are incommensurate with their abilities. Awards for scholastic achievement and other forms of social recognition have frequently been used to stimulate the student's motivation for higher education, but controlled studies of the effectiveness of social recognition as an incentive have been rare.

Previous studies have suggested that giving students social recognition for outstanding performances on aptitude tests tends to increase their motivation to attend college and their success in obtaining scholarship aid. Thistlethwaite (1958) estimates that, of the near winners receiving recognition in the second Merit competition, only about 3% did not enroll in college—a dropout rate which is less than half of the most conservative estimate (Phearman, 1949) made before the initiation of the Merit Scholarship program. It was hypothesized that this extremely low dropout rate was a consequence of the public recognition which the near winners received in the Merit program. Holland and Stalnaker (1957) report that two-thirds of the finalists not receiving Merit Scholarships in the first Merit competition managed to obtain scholarships from other sources and that half of these students attributed their success in part to the recognition they had received.

To obtain a more reliable and comprehensive estimate of the effects of recognition, the present *ex post facto* study compares two groups of talented students receiving different amounts of public recognition. The two groups differ, then, in the amount of social support they received for academic achievement; it is also possible that they differ in their self-evaluations, as a result of learning how they rank relative to other talented students. Since both groups received some degree of recognition, these results tend to give a conservative estimate of the effects. A comparison of more disparate groups would probably reveal larger differences. But even though the magnitude of effects is probably underestimated, some of the student behaviors influenced by public recognition are revealed by the present study.

METHOD

The first group (Group A) was selected from the 5126 students (80% of the Certificate of Merit winners in the 1957 program) who replied to a survey sent to all C of M winners. Each member of Group A had received a C of M attesting to his "high potential for college achievement" as demonstrated by a "distinguished performance on the nationwide selection tests for Merit Scholarships" and had had his name published in a booklet which was distributed to all accredited junior and senior colleges, to universities, and to a large number of other scholarship-granting agencies. In addition, these students were frequently publicized by press serv-

[1] The writer is indebted to Laura Kent for her helpful editorial comments on drafts of this paper. This study was partially supported by the National Science Foundation and the Old Dominion Foundation.

ices and acclaimed at high school assemblies. The second group (Group B) was drawn from 2848 of the second-highest-scoring 7500 students, who had received a letter of commendation and somewhat less recognition in the press. The pool from which Group B was drawn represents the 80% replying to a survey sent to a 47% random sample of commended students. No attempt was made to announce the names of the commended students to colleges and universities or to other scholarship donors. A count of press clippings shows that news items totalled approximately 2600 for C of M winners and 1100 for commended students—a ratio of approximately two and one-half to one.[3]

Matching Variables

So that the two groups would be comparable, C of M winners were matched with commended students on the following variables: (*a*) sex, (*b*) verbal aptitude score (the College Entrance Examination Board Scholarship Qualifying Test), (*c*) geographical region, and (*d*) father's occupational class. The majority of the responses to the question about father's occupation were assigned to classes on the basis of fixed-alternative choices; the remaining "free" responses were coded using the Minnesota Scale of Paternal Occupations (1955). From Holland's comparison of father's occupational class (as defined by the Minnesota Scale) with family income as reported by 800 finalists in the 1956 Merit program (1958), it is estimated that the tetrachoric correlation between these two variables is .64; thus, father's occupational class is a crude index of the family resources available for sending the student to college. In addition, a rough matching was made on mathematical aptitude (SQT-M). The use of these matching variables resulted in 1302 matched pairs, each pair made up of a C of M winner (Group A) and a commended student (Group B). There were no significant differences between the two groups with respect to SQT-V, geographical origin, or father's occupation. Sex distributions were identical. However, on the mathematical subscale of the SQT, Group A has a mean which was 1.7 units greater than the mean of Group B—a difference small in magnitude but nevertheless statistically significant ($p < .01$). Since both members of a pair were identical with respect to the four primary control variables but differed in the amount of public recognition they had received, it was possible to study the effects that such recognition has on students.

Measurement of Attitudes

To provide a measure of the impact of recognition upon attitudes toward intellectualism, the following eight items were included in the surveys (each response in the keyed direction was weighted one):

1. College teachers complain a lot about their pay, but it seems to me they get as much as they deserve. Disagree.
2. The work of theoreticians and scholars should be subsidized even though it may lack practical value. Agree.
3. Science and philosophy have their place, but there are probably many important things that can never be understood by the human mind. Disagree.
4. I spend more of my free time reading than I do watching television. Yes.
5. Most intellectuals are extremely competent persons and would excel in almost any job they undertake. Agree.
6. I am annoyed by writers who go out of their way to use strange and unusual words. No.
7. Generally the man of ideas is more important to society than the man of action. Agree.
8. I would rather write a fine book than be an important public figure. Agree.

[3] Commended students received their letters of commendation four to six weeks before the C of M winners received their certificates. Thus the commended students had the advantage of prior recognition in applying to other donors for scholarship aid.

Results

Attitudes Toward Intellectualism

One of the aims of the nationwide talent search annually conducted by the National Merit Scholarship Corporation is to call attention to intellectual attainment and, thereby, to stimulate promising young scholars and scientists to seek further specialized training. It is reasonable therefore to regard differences in attitude toward scientific or scholarly careers as one criterion of the effectiveness of the Certificate of Merit awards. Students in Group A tend to have higher intellectualism scores than those in Group B ($p < .01$). Further, recognition has statistically significant effects upon the attitudes of girls, students whose fathers are not employed in professional or managerial occupations, and students with relatively high verbal aptitudes. Boys and students whose fathers are in professional or semiprofessional occupations tend to be influenced in the same manner, although these influences are slightly smaller. The crucial question is, of course, whether these changes in attitudes are paralleled by similar changes in career plans.

TABLE 1

Effects of Social Recognition Upon Attitudes toward Intellectualism

Classification[a]	No. of pairs	Percentage of AB pairs in which A had higher intellectualism score	p[b]
High SQT-V	511	55.6	<.02
Low SQT-V	468	53.2	—
Boys	643	53.8	<.06
Girls	336	55.7	<.04
Professional or semiprofessional father	644	53.7	<.06
Nonprofessional father	335	55.8	<.03
Total.....	979	54.4	<.01

[a] The mean SQT-V score for "high aptitude" members of Group A was 48.3; for matched Group B students, 48.2. Corresponding mean scores for "low aptitude" A and B students were 41.3 and 40.8, respectively. The term "low aptitude" (which designated AB pairs in which the B member's score fell below the Group B median) is not descriptive in an absolute sense, since those students had a mean score above the 96th percentile relative to estimated national norms for public high school seniors.

[b] Two-tailed test based upon tests of the significance of correlated proportions have been used in this, and in succeeding, tables to estimate probability values.

Career Plans

An examination of the vocational plans of students in Groups A and B corroborates the findings suggested by the preceding analysis. Public recognition tends not only to change the "verbal" attitudes of students—but also to change their career plans. Students in each group were told to "check the vocation you now plan to enter as your life work." A significantly greater proportion of the students in Group A report they are planning to be college teachers or scientific researchers. Table 2 shows that this effect, like the effect upon attitudes toward intellectualism, occurs primarily among girls and among students whose fathers are not employed in professional or managerial occupations.

Despite the similarity of the results in Tables 1 and 2, it appears that items calling for an expression of vocational plans measure different processes than items which call for the expression of more diffuse beliefs and preferences. For example, changes in attitudes occur primarily among students with exceptionally high verbal aptitudes (Table 1), whereas changes in vocational plans occur primarily among students with only moderately high verbal aptitudes (Table 2). The former, it may be assumed, are influenced by recognition to place a greater value upon intellectual achievement in the arts, humanities, and social sciences, but are not attracted to careers in scientific research. Students who have only moderately high verbal aptitudes, on the other hand, are influenced by recognition to go into science careers, even though their attitudes toward scholarly pursuits do not change. In short, it is hypothesized that the in-

TABLE 2

EFFECTS OF SOCIAL RECOGNITION UPON PLANS TO BE A COLLEGE TEACHER OR SCIENTIFIC RESEARCHER

Classification	No. of pairs[a]	Percentage planning to be a college teacher or scientific researcher		Percent. diff.	p[b]
		Group A	Group B		
High SQT-V	672	24.2	21.7	2.5	—
Low SQT-V	620	23.5	17.9	5.6	<.02
Boys	850	24.9	22.2	2.7	—
Girls	442	21.9	15.4	6.5	<.02
Professional or semiprofessional father	827	23.4	20.4	3.0	—
Nonprofessional father	465	24.7	18.7	6.0	<.03
Total	1290	24.0	19.9	4.1	<.02

[a] The number of pairs vary among the tables because of nonresponse to survey items.

[b] Two-tailed tests of the significance of differences between correlated proportions have been used in this, and in succeeding, tables to estimate probability values.

TABLE 3

EFFECTS OF RECOGNITION UPON PLANS TO SEEK PHD OR MD DEGREE

Classification	No. of pairs	Percentage planning to get the PhD or MD degree		Percent. diff.	p
		Group A	Group B		
High SQT-V	662	37.3	34.0	3.3	—
Low SQT-V	620	38.7	32.8	5.9	<.05
Boys	845	49.1	45.1	4.0	—
Girls	437	16.5	10.5	6.0	<.02
Professional or semiprofessional father	821	40.8	35.8	5.0	<.03
Nonprofessional father	461	33.0	29.1	3.9	—
Total	1282	38.0	33.4	4.6	<.01

tellectualism scale is primarily sensitive to an intellectual orientation toward the arts, humanities, and social sciences; whereas the analysis presented in Table 2 primarily reflects differences in orientation toward a scientific career.

Motivation to Seek Advanced Training

Similar trends were observed when the educational plans of Groups A and B were compared. A significantly greater proportion of Group A reported plans to seek the PhD or MD degree (Table 3). Slightly greater—though not significantly greater—effects were observed among girls, students with only moderately high verbal aptitudes, and among students whose fathers are in professional or semiprofessional occupations.

Success in Obtaining Scholarships From Other Sources

Public recognition increases significantly the likelihood that the student will receive scholarship aid. 60% of the students in Group A and 55% of the students in Group B hold freshman scholarships ($p < .02$). Table 4 shows that the effects of recognition upon success in obtaining assistance occur primarily among boys, stu-

TABLE 4

EFFECTS OF RECOGNITION UPON SUCCESS IN OBTAINING SCHOLARSHIP ASSISTANCE

Classification	No. of pairs	Percentage of students holding scholarships		Percent. diff.	p
		Group A	Group B		
High SQT-V	616	59.7	57.1	2.6	—
Low SQT-V	582	60.3	52.9	7.4	$<.01$
Boys	791	61.3	56.3	5.0	$<.05$
Girls	407	57.5	52.8	4.7	—
Professional or semiprofessional father	791	53.2	47.8	5.4	$<.04$
Nonprofessional father	407	73.2	69.3	3.9	—
Total	1198	60.0	55.1	4.9	$<.02$

dents with only moderately high verbal aptitudes, and students whose fathers are employed in professional or semiprofessional occupations.

It has already been pointed out that in applying to other donors for scholarships, Group B had the advantage of prior recognition. It seems likely that, if the two groups had received their different degrees of recognition simultaneously, there would have been a greater difference between them with respect to their success in winning scholarships.

Proportions Enrolling in College

Because almost all of these students are college-bound, the hypothesized effects of recognition upon the decision to go to college could not be adequately tested. Approximately 96% of each group enrolled in college. Of those who did not enroll, the majority deferred their college plans only temporarily, usually because of military service or lack of financial resources. Because of the extreme marginals, no appreciable differences were expected—and none were found—between Groups A and B in the proportions entering college immediately.

DISCUSSION

The major effects of social recognition for outstanding performances on aptitude tests are: (*a*) to make more favorable the recipient's attitudes toward intellectualism, (*b*) to motivate him to enter a college teaching or a scientific research career, (*c*) to stimulate him to seek advanced degrees, and (*d*) to increase the likelihood that he will obtain scholarship assistance in college. It might be expected that public recognition has the greatest effect upon the "late-bloomer"—the highly talented student who has not had a particularly distinguished record of achievement in high school. Such students are, of course, the very ones a talent search seeks to discover and influence; those students whose talents have been largely unrecognized by their peers and teachers should be more influenced by public recognition than the valedictorians or salutatorians of high school classes. To check this hypothesis, additional comparisons were made of the AB pairs after they had been sorted by the high school rank of the C of M winner (unfortunately high school ranks were not available on members of Group B, so that these comparisons are less rigorous than desired). As expected, recognition produced the greatest increments in favorableness toward intellectualism among the students who had high verbal aptitudes (SQT-V scores above the group's median) but relatively undistinguished records of high school achievement (percentile ranks below the group's median). On the other hand, recognition

did not have an exceptionally great influence upon the "late-bloomer's" career plans or upon plans to seek the PhD or MD degrees. It may be that such students, being unaccustomed to evaluating themselves as persons of outstanding talent, need time to rethink their vocational plans. It is possible that marked "sleeper effects" will be found among these students—that is, the effects of public recognition upon the "late-bloomer's" career plans may not be fully manifest on an immediate after-test, but only after a year or more in college.[8]

In interpreting these results, one point should be kept in mind: both Group A and Group B received considerable public recognition. Although the two groups differed in type and amount of recognition received, differences were not as marked as those which would be found if "recognized" and "not recognized" groups were compared. Consequently, the magnitude of effects is probably less than that which would be observed in the ideal experimental comparison. The two groups are not perfectly matched, it is true, in mathematical aptitude, but the difference between them is very small. There is also the possibility—present in all *ex post facto* designs—that the observed effects are influenced by extraneous factors. On the other hand, the literature suggests that the matching control variables used account for most of the predictable variance in college-going behaviors (Berdie, 1954; Educational Testing Service, 1957; Hollingshead, 1952). On the bases of our present knowledge, therefore, we may conclude that all obvious extraneous variables have been reasonably well controlled, and that the results probably err in that they underestimate the effects of public recognition.

Summary

An *ex post facto* comparison was made of the educational motivation of two groups of talented students receiving different amounts of social recognition for their performances on college aptitude tests in a nationwide scholarship competition. Increased recognition was observed to increase the favorableness of attitudes toward intellectualism, the number of students planning to seek the PhD or MD degree, and the number planning to become college teachers or scientific researchers. The latter effect was observed particularly among students whose fathers are employed in nonprofessional occupations and among students who have only moderately high verbal aptitudes. In addition, increased recognition was observed to increase the likelihood that the students will obtain a scholarship from his college or from some other scholarship granting agencies.

[8] The present results are based upon surveys made approximately six months after graduation from high school.

REFERENCES

Berdie, R. F. *After high school—What?* Minneapolis: Univer. Minnesota Press, 1954.

Educational Testing Service. *Background factors relating to college plans and college enrollment among public high school students.* Princeton: Author, 1957.

Holland, J. L., & Stalnaker, J. M. An honorary scholastic award. *J. higher Educ.*, 1957, **28,** 361–368.

Holland J. L. A note on the reliability and validity of the Minnesota Scale for Paternal Occupations as an estimate of family economic status. *J. appl. Psychol.*, 1958, **42,** 195–196.

Hollingshead, B. S. *Who should go to college.* New York: Columbia Univer. Press, 1952.

Phearman, L. I. Comparisons of high-school graduates who go to college with those who do not. *J. educ. Psychol.*, 1949, **40,** 405–414.

Thistlethwaite, D. L. The conservation of intellectual resources. *Science*, 1958, **128,** 822–826.

University of Minnesota, Institute of Child Welfare. *The Minnesota Scale for Paternal Occupations.* Minneapolis, Minn.: Author, 1955.

Received October 22, 1958.

JOURNAL OF EDUCATIONAL PSYCHOLOGY
Vol. 50, No. 3, 1959

INTERCORRELATIONS AMONG VARIOUS INTELLIGENCE, ACHIEVEMENT, AND SOCIAL CLASS SCORES

LOTUS M. KNIEF

Wartburg College

AND JAMES B. STROUD

State University of Iowa

This investigation was planned, first, to provide additional data on the social class or culture bias issue in intelligence testing and, second, to ascertain interrelationships among certain relatively new intelligence tests and scholastic achievement. Scores on the following measures were intercorrelated: (*a*) Lorge-Thorndike Intelligence Tests, Verbal; (*b*) Lorge-Thorndike Intelligence Tests, Nonverbal (Level 3 in both instances); (*c*) Davis-Eells Games (Elementary Level); (*d*) Raven's Progressive Matrices; (*e*) Iowa Tests of Basic Skills; and (*f*) the Warner Index of Status Characteristics.

PROCEDURE

All of these tests except the Progressive Matrices were administered to a sample of 344 fourth grade pupils in a Midwestern city of about 80,000 general population. The basic sample consisted of all the pupils in this grade present at the time the tests were administered in 6 of the 18 elementary schools. The 18 schools were ranked upon the basis of mean scores achieved on the Iowa Tests of Basic Skills administered in a city-wide testing program. From this ranking every third school was selected and incorporated in the sample. The Progressive Matrices test was administered the following year to the pupils, now fifth grade, in every other fourth grade class in the six schools used originally. This procedure yielded only 164 pupils tested by the Progressive Matrices who had also taken all the other tests the year previously. The Iowa Tests of Basic Skills were administered in January 1957. The Davis-Eells test and the verbal and nonverbal tests of the Lorge-Thorndike battery were administered two or three weeks later; the Progressive Matrices, in the spring of 1958.

A modified Warner scale based upon occupation, house type, and dwelling area, weighted 5, 4, and 3, respectively, was used to assign social class ratings. Fathers' occupations were obtained from the school records. In 54 per cent of the cases the employers were interviewed in order to obtain a more precise job description. In assigning house-type and dwelling-area values the writers were assisted greatly by the City Assessor of the city involved. Pupils in the sample were rated on a 7-point scale on each of the three categories used, the assigned ratings were multiplied by the weights indicated above, and the products summed to provide a total ISC score value for each pupil in the sample. According to Warner estimates the following *n*'s would be expected for a sample of the size here used: 11, 34, 96, 117, and 86, for UU and LU (combined), UM, LM, UL, and LL classes, respectively. The obtained values were 11, 34, 100, 117, and 82.

In order to describe the sample further and to supply data on the behavior of the tests used, means and standard deviations are here given:

VARIABLE	MEAN	*SD*
Warner ISC	50.9	14.8
Davis-Eells IQ	104.5	15.4
L-T Verbal IQ	106.5	13.9
L-T Nonverbal IQ	108.8	13.7
ITBS Composite[1]	45.1	10.4

[1] Iowa Tests of Basic Skills scores are grade equivalent scores. 45.1 should be read as 5.1 months into the fourth grade.

Results

The zero order correlations among ISC, the intelligence scores, and ITBS Composite are shown in Table 1. As noted, the Progressive Matrices test was not administered at the same time as the other tests nor was it administered to all of the original sample. As supplemental data of some interest, correlations between raw scores on this test and scores on the other tests employed in this investigation were obtained and are reported in Table 1.

It may be of some interest to report the correlations between each of the intelligence scores and the various ITBS subtest scores. This is done in Table 2.

TABLE 1
Intercorrelations Among ISC, Intelligence Measures, and ITBS Composite Scores
(N = 344)

Variable	1	2	3	4	5	6
1. ISC		.309	.304	.323	.179	.340
2. D-E IQ			.571	.645	.405	.568
3. L-T V IQ				.709	.437	.839
4. L-T NV IQ					.521	.683
5. Prog. Matrices[a]						.450
6. ITBS Composite						

[a] Test 5, N = 164.

TABLE 2
Correlations between ISC, the Intelligence Measures, and Various ITBS Subtests
(N = 344)

	ITBS Tests				
	Reading Compre.	Vocabulary	Work-Study	Language	Arithmetic
ISC	.255	.323	.290	.365	.288
D-E IQ	.516	.561	.532	.446	.515
L-T V IQ	.776	.790	.752	.726	.739
L-T NV IQ	.580	.608	.679	.576	.663
Prog. Matrices[a]	.359	.438	.424	.374	.449

[a] N = 164.

TABLE 3
Partial Correlation Coefficients Among ISC, Intelligence, and ITBS Composite Scores
(N = 344)

Variables[a]	Partial Coefficient	Zero Order r's
51.4	.518*	51 = .568
52.4	.821*	52 = .839
53.4	.644*	53 = .683
54.1	.211*	54 = .340
54.2	.163*	.340
54.3	.173*	.340

[a] 5 = ITBS Composite
1 = D-E IQ
2 = L-T V IQ
3 = L-T NV IQ
4 = ISC
* Significant at .01.

Further analysis of the data was attempted by means of partial and multiple correlation procedures. Correlations were computed (a) between scores on each of the original three intelligence tests and ITBS Composite, with ISC scores partialled out, and (b) between ISC and ITBS Composite scores, with scores on each of the original three intelligence tests partialled out in turn. The coefficients are shown in Table 3, together with zero order correlations between the first two variables.

In the multiple correlation analysis Rs were computed between ITBS scores and various combinations of intelligence scores, taking two tests at a time. These are presented in Table 4, together with the zero order correlations between ITBS and the other tests involved, for the reader's convenience.

As a further step multiple regression analysis was performed with ITBS Composite scores as the dependent variable and ISC and the original three intelligence scores as independent variables. Beta weights were determined and tested for significance, regression equations developed, and multiple Rs computed. The results are reported in Table 5. The values

given above each column represent the *R* between ITBS and the other tests employed in that column.

By way of a brief discussion, attention is called to some of the more pertinent results obtained. All of the intelligence tests used and the ITBS Composite scores correlated significantly with social status (ISC scores), and, with the exception of the Progressive Matrices, to approximately the same extent. The partialling out of the various intelligence scores resulted in lower, but still significant, correlations between ISC and ITBS. The partialling out procedure tended to lower the correlations between ITBS and both the Davis-Eells and L-T nonverbal scores more than that between L-T verbal and ITBS. However, this tendency is more apparent than real. When the *r*'s involved are transformed into *z* values, the relationships between ITBS and the two kinds of intelligence scores (verbal and nonverbal) appear to be reduced about equally.

When, in multiple correlation, ISC scores were combined with the various intelligence scores, there was a slight tendency for this procedure to increase the *R*s between Davis-Eells and L-T nonverbal scores and ITBS more than the *R* between L-T verbal and ITBS. However, again, when *z* transformations are made, the differences are negligible. Our reasoning was that, if the Davis-Eells and L-T nonverbal tests are influenced less by cultural status than the L-T verbal test, combining ISC scores with those of the two former tests should raise their correlation with ITBS more than combining them with the verbal scores, since the latter, by some arguments, already contain social class loadings. This did not prove to be the case. It should be noted that the authors of the verbal test in question were cognizant of the culture bias factor and went to considerable length to minimize it.

Perhaps the most important result from the multiple correlation analysis is the light it sheds upon the behavior of the tests involved. The zero order *r*'s showed that the L-T verbal test gave the best prediction of ITBS scores, followed in order by L-T nonverbal and Davis-Eells. The multiple correlation analysis showed that L-T verbal alone correlated with ITBS about as well as did the entire battery of tests when combined in multiple correlation analysis. It may be noted that the deletion of L-T verbal scores from the regression analysis increased both the significance and magnitude of Davis-Eells Beta weight.

Progressive Matrices correlated to a smaller degree with ITBS than did any of the other intelligence tests. It may also be seen that the L-T verbal test correlated

TABLE 4

MULTIPLE CORRELATION COEFFICIENTS, THREE-VARIABLE COMBINATIONS

(N = 344)

Combination[a]	*R*	Zero Order *r*'s
5.23	.848	51 = .568
5.12	.846	52 = .839
5.24	.844	53 = .683
5.13	.703	54 = .340
5.34	.695	
5.14	.594	

[a] 5 = ITBS Composite
1 = D-E IQ
2 = L-T V IQ
3 = L-T NV IQ
4 = ISC

TABLE 5

MULTIPLE CORRELATION COEFFICIENTS AND BETA WEIGHTS FOR REGRESSION ANALYSES

(N = 344)

Multiple *R*s:	ITBS Composite Scores				
	.853	.851	.851	.849	.711
ISC	.067**	.074**		.076**	.111*
D-E IQ	.073		.083**	.117*	.200*
L-T V IQ	.687*	.702*	.694*	.749*	
L-T NV IQ	.127*	.161*	.137*		.518*

* Significant at .01.
** Significant at .05.

with every one of the ITBS subtests to a greater degree than did any of the other intelligence tests. The L-T nonverbal test correlated with every one of these subtests to a greater degree than did the Davis-Eells test. Moreover the latter correlated with every one of these subtests to a greater degree than did the Progressive Matrices.

The analyses reported in this paper give little justification for the use of the Davis-Eells Games, L-T nonverbal test, and the Progressive Matrices in conjunction with the L-T verbal test for *general prediction purposes*. This in no sense denies their usefulness in individual diagnosis. In need of further exploration in particular is the significance of extreme discrepancies in performance on verbal and nonverbal intelligence tests.

Received October 28, 1958.

JOURNAL OF EDUCATIONAL PSYCHOLOGY
Vol. 50, No. 3, 1959

SOME BEHAVIORS OF ELEMENTARY SCHOOL CHILDREN RELATED TO CLASSROOM ACTIVITIES AND SUBJECT AREAS[1]

SURANG KOWATRAKUL
Chulalongkorn University

In the field of educational psychology numerous questions concerning teaching-learning interaction have been awaiting empirical investigation. Many psychologists agree that empirical studies in the classroom itself will prove useful in a variety of ways. But this type of study when rigorously performed is rather difficult to accomplish because of several factors within the classroom situation. Furthermore, several studies (American Council on Education, 1945; Anderson, 1937) have indicated that teachers need new techniques and methods for gathering information to aid them in obtaining new insights into their pupils' behavior.

Educational psychology, therefore, would seem to need reliable instruments or methods for measuring students' behaviors in the classroom as well as reliable data which may form a solid foundation for the formulation of new concepts and the development of new theory. The aim of this study was to develop an instrument for systematic observation of pupils' behavior in what Sears has called a "naturalistic settings" situation (1957) and to make an empirical survey of the relationships of six categories of student behavior manifestations to three classroom activities and four subject areas.

[1] Based on a dissertation done at the Laboratory of Human Development and submitted to Stanford University as part of PhD requirements. Grateful acknowledgement is made to Pauline S. Sears, Quinn McNemar, and Frederick J. McDonald for guidance and to Pra Chuang Kashetra, Thai Cultural Counselor, for securing financial aid for the research.

METHOD

Definition of Behavior Categories

Intent on Ongoing Work (IW) or discussion, *task oriented*, performing assigned work. This category includes active work in assigned areas and also collecting, getting out, organizing, and putting away materials necessary for work. The key phrase is *task oriented*.

Social Work Oriented (SW): This category includes any social remark, interchange, or action which is work oriented. For example, a child may initiate social contacts in relation to his own work, he may respond to another child's work, he may explain a task to a peer, he may note another's progress or show his own progress, give or get suggestions in regard to work, or he may initiate interaction with a teacher in regard to a work problem. This category is reserved for the expression of social needs in relation to work tasks.

Social-Friendly (SF): This category includes any social remark, interchange, or action. Within this category fall mutual horseplay, friendly conversation and gestures, and any funny faces or movements to attract others' attention. The key words in this category are *pure social* interchange without task orientation to ongoing work.

Momentary Withdrawal (MW): This category includes the behavior which occurs when an individual momentarily ceases his present task and wanders or looks around, watches others without apparent reason; or withdraws from class activity, appears unaware of the group, stares at a particular spot or into space, assumes an impassive expression and is

not, apparently, thinking about his work or engaging in social activity.

Intent on Work in Another Academic Area (WOA): Tallies fall in this category when a student is doing some other academic work than that which has been assigned for the present period.

Intent on Work in Nonacademic Area (WNA): This category is checked when a student is intent on activity of his own which is unrelated to academic work. This category also includes all complex play or doodling activity which is purposive or directed toward a consciously defined goal. For example, the student is making a paper airplane, or cleaning his desk, or arranging a string or chain or paper clips into a certain figure.

Definition of Classroom Activities

Independent Seat Work: In this category, students are engaged in work individually. Usually a student does an assigned job which has a clearly defined goal.

Watching and Listening: This category of classroom situation is employed when students are supposed to play a passive role. Examples: listening to the teacher, watching a demonstration, watching an educational program on TV.

Discussion may be defined as symbolic interaction among three or more persons. Such interaction may be structured by a participating formal leader, student or teacher; or it may take place within a structure prearranged by the teacher with a leader appointed by her; or it may be at a purely informal level following the bent of the participants. Discussion may be divided into small and large groups.

Definition of Subject Areas

Language is defined as any kind of work directed towards acquiring a proficiency in a manipulation of symbols (oral or written expression), their structuring, ordering, construction, and pronunciation. Examples include spelling, reading recreational books, and writing or giving an oral report where emphasis is on successful communication rather than content.

Arithmetic is defined as any work concerned with numbers and their relationships.

Social Studies consists of work in human relations, civics, current events, history, news, and world events.

Science consists of any study concerned with the principles of nature, facts of natural phenomena, or methods of the biological, physical, and social sciences.

Sample

56 children in Grades 5 and 6 of a small elementary school were observed. This number of *S*s includes 15 boys in Grade 5 and 18 in Grade 6, and 10 girls in the former class and 13 in the latter.[2]

Method of Gathering Data

This study employed the point-time sampling technique in which an observer observes an *S* long enough to record one behavior, then passes on to the next *S* immediately. The behavior of the next *S* must be independent of that of the first in order to ensure independent behavior scores. Observations in this fashion continue until one "behavior point" of each individual in the group has been recorded. Then a new round of observation begins. Observation experience indicated that each round took about four to five minutes in the sixth grade class and about three to four minutes in the fifth grade class.

For efficiency and preciseness in recording, observation record sheets were prepared prior to the time of gathering data. The observer also described in detail the nature of the classroom activity at the time of observation. Data were recorded

[2] Although sex differences are likely for some of the six behaviors defined above, analyses separately for the sexes were not contemplated because of consequent small *N*s and because the emphasis is on behaviors of a class as a whole. The results, therefore, cannot safely be generalized to sex segregated classes.

only when classes were engaged in the particular activities defined above. No recordings were made during intervals of transition from one activity to the next. Sometimes classroom activities were diversified in which case behavior frequencies were not recorded. Data were not collected when a substitute teacher replaced the regular one.

Further, because this study was a "naturalistic settings" type, the observer tried not to interfere with the plan of the class. The observer also attempted to reduce any halo effect which might arise from her subjective impressions of the pupils. During the time of collecting data she tried to avoid contact with individual pupils.

Table 1 presents the mean number of time samples per student under each classroom activity and subject area in each grade. The total number of time samples was 45,526.

TABLE 1

MEAN NUMBER (PER STUDENT) OF TIME SAMPLES FOR THREE CLASSROOM ACTIVITIES AND FOUR SUBJECT AREAS

(*Ns*: Grade 6, 31; Grade 5, 25)

Classroom activity	Grade	Subject Area			
		Science	Soc. Stud.	Arith.	Lang.
Independent Seat Work	6	65	60	66	84
	5	82	68	63	89
Watching-Listening	6	118	77	61	62
	5	62	62	55	61
Discussion	6	65	65	64	51
	5	59	61	62	58

Scoring and Reliability

For reasons beyond the control of the experimenter the number of time samples (or observations) was not constant for all students, hence it was necessary to use a percentage scoring system. Thus, e.g., a score of a student for the IW category is the number of behaviors judged as IW relative to the total number of observations on that student under a specified condition (such as Arithmetic-Discussion). For each student there were 12 scores (corresponding to the 12 combinations of conditions) for each of the six behavior categories.

Score reliability was ascertained by correlating percentage scores for odd-numbered vs. even-numbered time samples. Since the 30 computed Spearman-Brown coefficients (Table 2) were generally high,

TABLE 2

SPEARMAN-BROWN RELIABILITIES OF BEHAVIOR SCORES

Classroom Activity	Subject Area	Behavior Categories					
		IW	SW	SF	MW	WOA	WNA
Grade 5							
Watching-Listening	Science	.982	.986	.956	.970	.894	.985
Independent Seat Work	Language	.984	.953	.975	.960	.984	.960
Discussion	Social Studies	.941	.928	.894	.881	.932	.943
Grade 6							
Independent Seat Work	Arithmetic	.957	.942	.737	.969	.953	.953
Discussion	Science	.943	.994	.805	.905	.875	.988

it seemed unnecessary to determine all of the possible 144 reliability coefficients. Many of the computed reliabilities are sufficiently high to suggest that adequate reliability could be attained with fewer observations.

Perhaps, in the light of the given reliabilities, it is not necessary to report that the preliminary question of observer reliability received a satisfactory answer. Data were collected over a period of five days by a trained observer (M. H. Pintler of the Laboratory of Human Development) and the writer. These two observers simultaneously fixated upon the same student, with identical observation times ensured by the use of signals. Following the signal, the behavior of a student was immediately observed, categorized, and recorded independently by the two observers. For a total of 1,689 time samples, with data combined for different classroom activities and subject areas, there was 94% agreement in the categorizing. This figure is sufficiently high to indicate that the data collected by a single observer could be replicated by another observer.

TABLE 3

MEANS FOR INTENT ON ONGOING WORK (IW) UNDER THREE CLASSROOM ACTIVITIES AND FOUR SUBJECT AREAS

(Grade 6 top half; Grade 5 bottom half)

	Science	Social Studies	Arithmetic	Language	Mean of Means
Independent Seat Work	76.07	82.64	78.13	74.16	77.75
Watching-Listening	73.55	71.77	73.35	76.39	73.77
Discussion	52.42	53.45	52.71	56.94	53.88
Mean of Means	67.34	69.29	68.06	69.16	68.47
Independent Seat Work	82.92	81.48	79.08	77.68	80.29
Watching-Listening	82.12	84.80	74.40	79.28	80.15
Discussion	52.56	56.00	46.88	53.24	52.12
Mean of Means	72.47	74.09	66.79	70.07	70.85

ANALYSES OF DATA AND FINDINGS

It was stated in the presentation of the method of gathering data that the frequencies for the occurrence of each type of behavior (Intent on Ongoing Work, Social Work, Social-Friendly, Momentary Withdrawal, Intent on Work in Another Academic Area, Intent on Work in Nonacademic Area) of each child were collected under each condition of three classroom activities (Independent Seat Work, Watching-Listening, and Discussion) and four subject areas (Science, Social Studies, Arithmetic, and Language). The design is such that tests of significance involve a three-way analysis of variance, separately for each grade, with rows representing individuals, columns representing the four subject areas, and blocks representing three kinds of classroom activities. This type of analysis was made for each dependent variable except Intent on Work in Another Academic Area for which the distributions were highly skewed.

With I standing for Individuals, S for Subject Areas, and A for Classroom Activity, the main effect for S was tested against the I × S interaction, the A effect against the I × A interaction, and the S × A interaction against the I × S × A interaction. Because of space limitations the usual analysis of variance tables, 10 in number, will not be presented. Instead, tables of means will be given, with the results for Grades 5 and 6 in the top and bottom halves, respectively. The statistical significance of the comparisons will next be briefly summarized.

For Intent on Ongoing Work (see Table 3) the Activities (A) effect is significant beyond the .0005 level for both grade groups, with markedly lower IW during Discussion. The Subject Area (S) effect is significant (.0005 level) for Grade 5, but insignificant for Grade 6. The A × S in-

TABLE 4

MEANS FOR SOCIAL WORK (SW)

	Science	Social Studies	Arithmetic	Language	Mean of Means
Independent Seat Work	8.87	6.94	7.29	5.55	7.16
Watching-Listening	7.36	3.68	9.19	7.19	6.86
Discussion	22.16	12.74	28.61	23.39	21.73
Mean of Means	12.80	7.78	15.03	12.04	11.91
Independent Seat Work	9.20	10.44	10.20	9.84	9.92
Watching-Listening	6.68	4.88	17.52	9.92	9.75
Discussion	31.24	26.32	43.36	32.00	33.23
Mean of Means	15.71	13.88	23.69	17.25	17.63

TABLE 5

MEANS FOR SOCIAL-FRIENDLY (SF)

	Science	Social Studies	Arithmetic	Language	Mean of Means
Independent Seat Work	6.13	4.71	5.55	8.81	6.30
Watching-Listening	5.16	5.74	5.13	4.10	5.03
Discussion	6.52	10.97	5.07	5.35	6.98
Mean of Means	5.94	7.14	5.25	6.09	6.10
Independent Seat Work	2.60	2.84	2.28	4.20	2.98
Watching-Listening	2.84	2.44	3.68	1.72	2.67
Discussion	3.48	4.76	5.32	2.40	3.99
Mean of Means	2.97	3.35	3.76	2.77	3.21

teraction reached the .005 level for Grade 6, but was not significant for Grade 5.

For the SW variable (Table 4) both main effects and their interaction are highly significant (beyond .0005 level) for both grade groups. High SW occurs consistently (despite the A × S interaction) during Discussion, but the high SW for Arithmetic compared to low for Social Studies needs to be qualified because of the A × S interaction.

The Social-Friendly behavior (Table 5) yielded significant A × S interaction (beyond the .001 level for both grades), with insignificant S effect for Grade 5, and only .05 level of significance for Grade 6. The A effect was borderline in significance (.03 level for both grades). The main effects are overshadowed by the interactive effect.

The only significant effect for MW (Table 6) is the A × S interaction for Grade 6 (.0005 level).

TABLE 6

MEANS FOR MOMENTARY WITHDRAWAL (MW)

	Science	Social Studies	Arithmetic	Language	Mean of Means
Independent Seat Work	4.77	3.07	4.68	6.71	4.81
Watching-Listening	5.13	6.48	4.45	3.90	4.99
Discussion	4.07	6.71	4.13	4.58	4.87
Mean of Means	4.66	5.42	4.42	5.06	4.89
Independent Seat Work	3.04	2.00	3.56	3.12	2.93
Watching-Listening	2.28	2.20	1.60	2.52	2.20
Discussion	2.72	3.56	2.24	4.60	3.28
Mean of Means	2.68	2.59	2.53	3.41	2.80

For behavior defined as Work in Nonacademic Area (Table 7), all three effects for the sixth grade group were significant beyond the .001 level as were also the A and the A × S effects for Grade 5, with the S effect for the latter group reaching only the .03 level.

As mentioned earlier the distributions of the scores of Intent on Work in Another Academic Area (WOA) were highly skewed. Cochran's *Q* formula (McNemar,

TABLE 7

MEANS FOR INTENT ON WORK IN NONACADEMIC AREA (WNA)

	Science	Social Studies	Arithmetic	Language	Mean of Means
Independent Seat Work	3.10	2.36	2.77	4.03	3.07
Watching-Listening	7.16	8.13	5.87	5.45	6.65
Discussion	11.45	13.06	8.06	6.87	9.86
Mean of Means	7.24	7.85	5.57	5.45	6.53
Independent Seat Work	1.16	2.12	2.64	3.68	2.40
Watching-Listening	4.40	3.96	2.24	3.72	3.58
Discussion	8.04	7.40	2.24	5.64	5.83
Mean of Means	4.53	4.49	2.37	4.35	3.94

p. 232) was employed to analyze these scores. Cochran's technique as here used is an extension of the nonparametric "median" test to the situation involving correlated scores. This method does not permit a test of the A × S interaction. The S effect was significant at the .01 and .02 levels for the sixth and fifth grades, respectively. The A effect was significant at the .001 level for Grade 6, but insignificant for Grade 5. The percentages exceeding the appropriate medians are given in Table 8,

TABLE 8

INTENT ON WORK IN ANOTHER ACADEMIC AREA (WOA)—PERCENTAGES EXCEEDING MEDIAN VALUES

Conditions	Grade	
	6	5
Science	55	44
Social Studies	65	64
Arithmetic	29	24
Language	52	56
Independent Seat Work	23	36
Watching-Listening	65	56
Discussion	55	52

from which it will be noted that there was less WOA during Arithmetic and during Independent Seat Work.

No comparisons were made for any of the behavior scores between the two grades because such differences as were obtained are possibly confounded with a teacher effect. Incidentally, the magnitudes of the over-all means in Tables 3–7 are indicative of the relative amounts of the several categories of behavior. Thus about 70% of the behavior during the observation times was Intent on Ongoing Work, with about 12% and 18% for the sixth and fifth grades, respectively, being Social Work. The percentages for the other categories are relatively small and homogeneous.

DISCUSSION AND INTERPRETATION

We may conveniently group the six students' behavior manifestations into two categories with respect to the kind of teacher response they evoke: teacher-approved and teacher-disapproved. In the first of these groups fall the behaviors which were called Intent on Ongoing Work and Social Work. These two types of behavior were reinforced by teachers who often gave verbal approval to students who manifested them. The second group, Teacher-Disapproved behavior, includes the Social-Friendly, Intent on Work in Another Academic Area, Intent on Work in Nonacademic Area, and, perhaps, Momentary Withdrawal. As indicated in the just previous paragraph, over 80% of the observed behavior fell in the teacher-approved category.

Effect of Classroom Activities

Does the particular nature of each classroom activity influence behavior? If one compares the means of the means of Intent on Ongoing Work behavior for classroom activities, it will be seen that the highest means are for Independent Seat Work and Watching-Listening for both classes. It would seem that Discussion activity is less conducive to this type of desirable or

teacher-approved behavior, whereas Discussion is favorable to Social Work behavior (Table 4). The high scores in Intent on Ongoing Work behavior in Independent Seat Work and Watching-Listening can be related to the nature of the activities themselves. In the activity called Independent Seat Work usually there were pressures for individually intensive work. For example, this activity involves tasks which must be completed within a certain time limit, research for individual reports, and creative writing—on which teachers base in part their evaluations. The high IW means for the Watching-Listening activity can be explained by the nature of this type of activity and tasks: usually students were required to take notes and listen carefully. The teacher often asked the class questions at the end of the activity. Further, Watching-Listening frequently aroused students' interest because of some novelty as exemplified by motion pictures, television productions, and guest speakers.

The relatively higher Social Work means for Discussion are related to the purpose of this activity which satisfies a social need for interpersonal interaction and exchange of ideas.

If we examine the means of the means (Table 7) of the category Intent on Work in Nonacademic Area for classroom activities, we see that in both classes the maximum values of the means of the means were in the Discussion activity. The results might seem to indicate the possibility that the Discussion has two characteristics: among students who have more self-confidence the discussion period may provide an opportunity to participate actively in talking and exchanging ideas; among more self-conscious students the discussion period is largely a time of watching and listening activity. Since these students were only listening, their hands were not engaged in taking notes; sometimes they drew pictures or figures.

The differences between the means of the means of Intent on Work in Nonacademic Area in Watching-Listening and Independent Seat Work activities were significant for both grades, with higher scores found in Watching-Listening activity. This finding again can be explained in terms of our previous analysis of the nature of the Watching-Listening activity, i.e., during the period students were not taking notes, they could engage in other activities such as playing with their pencils, erasers, or paper clips.

Any conjecture about the relatively small effect of Activity on the Social-Friendly behavior will need qualification because of the marked A × S interaction. It is not surprising that Independent Seat Work provides an opportunity for more Work in Another Academic Area.

Behavior Related to Subject Areas

For Grade 6, the over-all test for the main effect of Subject Area on the behavior Intent on Ongoing Work was insignificant, but in Grade 5 Subject Area had an effect. This contrast may reflect either teacher differences or a true grade effect.

For both grades the highest value of the mean of the means of Social Work behavior was for Arithmetic (Table 4), with the lowest in Social Studies. These results may derive in part from the relatively high degree of questioning both by the teachers and by students requesting help, during Arithmetic with fewer questions during Social Studies. It should be noted that it was not feasible to include in the research design the classroom activity, working in small groups, which was used much during Social Studies and which would be conducive to SW. Any further speculating about the effect of Subject Area on SW should consider the fact of A × S interaction.

Since the main effect of Subject Areas on Social-Friendly behavior was insignificant for Grade 5 and reached only the .05 level for Grade 6, and since the A × S

interaction was highly significant for both grades, any interpretation of the difference between the means of the means becomes precarious. Note in the top half of Table 5 that under Social Studies we have the highest mean (10.97) of the row for Discussion but the lowest (4.71) in the row for Independent Seat Work. Similarly, note the 4.20 and 1.72 in the bottom half of the table—Language leads to most SF in one, and least in another, Activity.

With regard to the differences in Intent on Work in Nonacademic Area associated with Subject Areas, it is interesting to note that the lowest value of the mean of the means for Grade 5 was under Arithmetic while in Grade 6 the means of the means were lowest for Arithmetic and Language. Once again, the presence of A × S interaction cannot be ignored when considering the S effect.

The effect of Subject area on WOA (see Table 8) indicates less WOA under Arithmetic. Apparently, Arithmetic was handled by these teachers so as to lead to less teacher-disapproved behavior.

Summary

This study was an investigation of the relationships of six behaviors manifested by 56 elementary school children (fifth and sixth graders) to three classroom activities and four subject areas. The findings demonstrated that classroom activities and subject areas and some of the behavior categories observed are significantly related.

From the findings of this study one may not legitimately infer that the behavior of all sixth and fifth graders will replicate the patterns uncovered here. This study did not have as its purpose any attempt to draw generalizations for a population but rather hoped to unfold some suggested relationships among classroom activity, subject area, and students' behavior manifestations which may on more exhaustive investigation prove useful in the formulation of a consistent theoretical framework.

A secondary purpose for this study was that of developing a reliable instrument for measurement of children's behavior which might be applied in classrooms. This was accomplished.

REFERENCES

American Council on Education, Division on Child Development and Teacher Personnel. *Helping teachers understand children.* Washington, D. C.: Author, 1945.

Anderson, H. H. Domination and integration and the social behavior of young children in an experimental play situation. *Genet. psychol. Monogr.*, 1937, **19**, 341–408.

McNemar, Q. *Psychological statistics.* New York: Wiley, 1955.

Sears, P. S. Problems in the investigation of achievement and self-esteem motivation. In M. R. Jones (Ed.), *Nebraska symposium on motivation.* Lincoln: Univer. Nebraska Press, 1957. Pp. 265–339.

Received November 22, 1958.

JOURNAL OF EDUCATIONAL PSYCHOLOGY
Vol. 50, No. 3, 1959

REFUSALS AND ILLEGIBILITIES IN THE SPELLING ERRORS OF MALADJUSTED CHILDREN[1]

HARRY H. L. KITANO
University of California, Los Angeles

Disability in subject matter areas, such as reading and spelling, have been a major target for research. For example, Gates and Russell (1940) have attempted to hypothesize emotional difficulties as possible causes for reading and spelling failures. Bower (1958) has shown that adjustment-class children (also known as emotionally disturbed and behavior-problem children) score significantly lower than regular-class children in reading, arithmetic, and other subject matter areas. The purpose of this paper is to study the quantitative and qualitative errors in spelling between regular- and adjustment-class children.

Logically, it appears that adjustment-class children should make more errors and different types of errors than their regular-class counterparts. Their overall poor adjustment in school (Bower, 1958), lower reading ability (Bower, 1958), and other related factors indicate poorer spelling achievement. These, plus their higher anxiety and rigidity (Kitano, 1958) and poorer ego controls (Kitano, 1958) suggest the following hypotheses: (*a*) that adjustment-class children will be poorer spellers than regular-class children; (*b*) that adjustment-class children will make different types of errors than regular-class children.

Definitions, Sample, and Delimitations

Adjustment classes are set up for behavior-problem children in the San Francisco Unified School District. The children are often referred to as the "emotionally disturbed"; the common underlying symptom is the inability to get along in a regular class assignment. The classes are set up for children who have normal intellectual potential.

The regular-class children in the study were drawn from one school in a lower-middle class area. This roughly corresponds to the socioeconomic status of the adjustment-class group as reported in a previous study (Kitano, 1958). The principal of this school describes her children as being from "fair to poor" in spelling achievement and reports the mean IQ to be around 95. This again roughly corresponds to the IQ level of the adjustment-class group (Kitano, 1958). Therefore, an attempt was made to match the groups in regard to socioeconomic status and IQ through this selection procedure.

The sample consisted of all the fourth, fifth, and sixth grade children from the regular-class group of one school and all the fourth, fifth, and sixth grade children in the adjustment classes. The number in each group was 88; the sample consisted of 176 children.

The spelling words were chosen from Gates' *A List of Spelling Difficulties in 3876 Words* (1937). Twenty words classified by Gates as being between the 4.6 and 4.9 grade levels were chosen as it was believed that these words would be familiar, yet difficult enough for fourth, fifth, and sixth grade children.

In a paper devoted to spelling errors, the delimitation of the method of classification of such errors is of utmost importance. Spache (1940) reviews the various classifications and suggests various criteria which must be met in order to analyze spelling errors. The method of classification used in

[1] The research was made possible through the valuable cooperation of Margaret Holland, Supervisor of Elementary Guidance; Clara Moore, Principal; and the adjustment class teachers of the San Francisco Unified School District.

this research is a combination of the Spache (1940) and Gates and Russell (1940) classifications, modified for the exploration of the hypotheses under investigation. The six types of errors and classifications are listed below:

Additions and insertions, such as sticke for stick
Omissions, such as fether for feather
Phonetic errors, such wate for wait
Substitutions and reversals
Words refused or not completed
Unrecognizable

Procedure

All testing was done in December of 1957. The writer gave the spelling test to the regular-class children; the individual adjustment-class teachers gave the test to their own classes.

Coding of the spelling errors was done by the experimenter.

A comparison of the type of errors was made by group, and statistical significance was tested through chi square.

Results

The summary of findings is presented in Table 1. The regular-class children made significantly more errors in additions, phonetics, and substitutions. The adjustment-class children made significantly more errors through refusals and unrecognizable spelling. There was no significant difference between the groups in omissions.

TABLE 1

Comparison and Significance of Total Errors by Groups in Additions, Omissions, Phonics, Substitutions, Words Refused, and Unrecognizable Spelling

	Adjustment-Class Errors	Regular-Class Errors	Chi Square	*P*
Additions	70	110	8.8	.02
Omissions	80	110	4.7	.10
Phonics	70	114	11.5	.01*
Substitutions	161	342	65.0	.01*
Words Refused	616	155	2756.4	.01*
Unrecognizable	312	155	52.9	.01*
Cases	88	88		
Spelling Words	20	20		

* Significant at .01 level.

The regular-class children scored significantly higher on the number of words correct. Of the total number of attempted and legible words, the adjustment-class children missed 52.7%; the regular-class children 17.6%.

Discussion

Causes of spelling errors have been classified into three broad categories by Cole (1936), and Gates and Russell (1940). The errors include (*a*) deficiencies within the pupil, (*b*) difficulties inherent in the English language, and (*c*) inappropriate methods of teaching.

The errors under the first two categories were explored in the present study. There was no attempt to control or analyze teaching methods.

It is possible to think of errors in additions, omissions, phonetics, and substitutions as being errors primarily due to the difficulty of the English language, or external errors. It is of interest to note that the regular-class children made more external errors than adjustment-class children.

The type *a* errors, or difficulties within the pupil, manifest themselves in errors due to "refusals" and "unrecognizable spelling." Something within the child makes it impossible for the child to try writing the word or to come up with something completely unrecognizable when he does write. It is suggested that the higher anxiety of the adjustment-class child (Kitano, 1958) is one reason for the "refusals" and the unrecognizable spelling. The findings appear to confirm an earlier study by the author (Kitano, 1958) in which adjustment-class children were found to have significantly higher anxiety and rigidity scores than regular-class children. Higher anxious and higher rigid children should be characterized by a refusal to try some-

thing new—in this case spelling words—while children with less anxiety and rigidity should feel relatively free to try.

In effect, the high anxiety blocks learning, the high rigidity channels the anxiety into a refusal to attempt or try words, and the rigidity and the refusal to try evidently creates more anxiety. This vicious circle in one specific subject matter area is no doubt repeated over and over again in school and is one probable cause for much of the discharge of anxiety by adjustment-class children through extremely hostile, aggressive behavior.

It is evident that the role of a teacher in an adjustment class is to somehow alleviate some of the excessive anxiety, perhaps by giving praise to those who try, no matter what the result, and to minimize the competitiveness of the normal spelling lesson. Success at a lower level, such as reduced spelling load, longer time intervals for learning, and the minimization of the formalized spelling test can be of possible value towards alleviating some of the anxiety producing situations.

REFERENCES

BOWER, E. M. *A process for early identification of emotionally disturbed children.* Sacramento, Calif.: California State Department of Education, 1958.

COLE, L. *The elementary school subjects.* New York: Rinehart, 1936.

GATES, A. I. *The psychology of reading and spelling with special reference to disability.* New York: Columbia Univer. Teachers Coll., 1922.

GATES, A. I. *A list of spelling difficulties in 3876 words.* New York: Columbia Univer. Teachers Coll., 1937.

GATES, A. I., & RUSSELL, D. *Diagnostic and remedial spelling manual.* New York: Columbia Univer. Teachers Coll., 1940.

KITANO, H. *Anxiety and rigidity in adjustment-class children.* Unpublished doctoral dissertation, Univer. of California, Berkeley, 1958.

SPACHE, G. A critical analysis of various methods of classifying spelling errors. *J. educ. Psychol.*, 1940, **31**, 111–134.

Received November 28, 1958.

JOURNAL OF EDUCATIONAL PSYCHOLOGY
Vol. 50, No. 3, 1959

PERSONALITY AND INDEPENDENT STUDY[1]

KATHRYN KOENIG AND W. J. McKEACHIE
University of Michigan

Does it make any difference how we teach? Despite a number of carefully executed studies on the comparative effectiveness of various teaching methods, there is little evidence to support the view that one teaching method is more effective than any other.

At the moment, our University of Michigan project is most vitally concerned with one explanation of the finding that there is little difference in the performance of students taught by one method as compared to other methods. Our assumption is this: "Any teaching method is effective only for certain students; when we compare teaching methods on group measures, the effects upon different students cancel each other out."

In this study we hypothesized that the highly independent students would prefer learning, perform better, and be more involved in an independent study situation. Likewise, we hypothesized that students with high need for affiliation would prefer, perform better, and be more involved in small group discussions than would other students. Finally, we hypothesized that students high in n Achievement would do well in independent study.

DESIGN OF THE EXPERIMENT

The experimental groups consisted of 124 students (89 females, 35 males) who were enrolled in a single lecture section of an elementary psychology course which met twice weekly. The students also met twice weekly in six discussion sections of about 20 students each.[2]

In order to test our hypotheses we manipulated the teaching methods in these sections experimentally so that during the semester each student would participate in small group discussions and independent study in lieu of the regular discussion classes. In order to achieve this, the regular discussion sections were discontinued for two weeks and each student in one half of the class wrote a paper independently, while students in the other half met in small groups. Later in the semester, the experimental procedure was repeated with the groups reversed so that each student participated in both variations of teaching method.

Students in the lecture group took the following measures of personality: The California Psychological Inventory, which gave scores on 18 traits; Stott's inventory, Every-day Life, which yielded three measures of self-reliance; and finally the Thematic Apperception Test, as modified by McClelland, which was scored for Need for Achievement, Need for Power, and Need for Affiliation. In addition, observers recorded the interaction in the small group discussions, using our modified Bales recording system.[3]

Measures of the students' preferences, involvement, and performance were administered after each variation of teaching and at the end of the semester. Students were asked to rate the importance of the teaching method and the effectiveness of the method. In addition the student indi-

[1] This study was part of a larger study supported by a grant from the Ford Fund for Advancement of Education. Substantial portions of this paper were presented in a symposium, "Experimental Studies in Learning Independently," at the 1958 APA meetings.

[2] The cooperation of the discussion leaders, Myron Braunstein, Nathan Brody, and Sarah Curtis is greatly appreciated.

[3] Richard Mann devised the categorization system used and trained our observers.

cated which method he preferred and how involved he was in that portion of the course. The instructors evaluated the students' performance in the small groups, on the paper, and in the total course.

Results and Discussion

Since previous studies had indicated that different results may occur for men and women, the analyses were done separately for each group. In general, the personality variables were used as independent variables; and the ratings, grades, and observation records were used as dependent measures.

The first analysis of the data was disappointing. The two major hypotheses were not supported. The measures of self-reliance or of n Affiliation were not related to satisfaction, performance, or involvement in the experimental groups.

When we looked at n Achievement, our results were a little more encouraging. Although there were no significant differences for men, high n Achievement women preferred the innovations (both small group and independent study sessions) to lectures. The middle n Achievement women, on the other hand, preferred lectures. Previous studies have indicated that middle n Achievement people are high in fear of failure and thus may wish to avoid the more threatening independent study and small group procedures because they demand new skills which they are not sure of. High n Achievement women, on the other hand, may find these innovations more challenging than the usual lecture sections.

Some of our peripheral results still puzzle us. For example we found some surprising sex differences. High n Power women participate less in the small groups than do low n Power. However, the direction of this relationship is reversed with men.

The trait of flexibility is even more puzzling. It is related to participation in the small groups. This CPI scale has a high negative correlation with the California F (Authoritarian) Scale, and it is not surprising that the flexible nonauthoritarian individual should participate freely in the rather permissive small groups. This is what we find for women. For men, how-

TABLE 1

Relationship Between Women's Need for Achievement and Teaching Method Preferences

Need Achievement (Women)	Preference for Method		
	Lecture	Small Group	Independent
High	2	8	5
Mid	16	4	3[a]
Low	7	7	6[b]

[a] Chi square = 12.6.
[b] Probability = .05.

TABLE 2

Relationship Between Women's Need for Power and Amount of Participation in Small Group Discussions

Need Power (Women)	Amount of Participation	
	High	Low
High	9	17[a]
Mid	7	16[b]
Low	14	7

[a] Chi square = 7.1.
[b] Probability = .05.

TABLE 3

A Comparison of the Participation of High Need Power Men and High Need Power Women in the Small Groups

High Need Power	Amount of Participation	
	High	Low
Men	10	2[a]
Women	10	23[b]

[a] Chi square = 10.0.
[b] Probability = .001.

ever, this scale predicts in the opposite direction. Highly flexible men are significantly *less* likely to participate than men low in flexibility.

Despite the number of negative results, this study provides some encouragement for continuing examination of individual differences with teaching methods. Our measures of need for power and of flexibility do predict participation in permissive small groups, albeit in odd ways. We once again found evidence that students who fear failure prefer familiar well-structured situations such as lectures, and we once again found sex differences even though we cannot explain them.

Some additional results maintain our feeling that programs of independent study need to take account of personality. Carleton College also participated in the Ford Fund studies of independent work.[4] 64 of their seniors took one of several independent study courses last fall semester. There were no control groups, but the students did take the CPI, Allport-Vernon, and a scale developed at Michigan to measure students' conception of the teacher's role.

While the Allport-Vernon and CPI failed to predict performance in independent study, our Michigan scale did. Students who think the instructor should be authoritarian tended to do poorly in independent study. However, this measure does not relate significantly to over-all grades in other college courses or to intelligence.

Summary and Conclusions

Each student of an introductory psychology course participated in small group discussions, independent study, and regular lecture - discussion sections. Personality data and measures of the students' preferences and performance under each method were collected. Neither the measures of self-reliance nor of n Affiliation were related to satisfaction, performance, or involvement in the small group discussions, or independent study. However, high n Achievement women preferred the two innovations to lectures, while middle n Achievement women preferred the lecture method. Additional findings were that high n Power women participated less in the small groups than low n Power women; whereas the direction of the relationship was reversed for men. Similarly, the nonauthoritarian women participated more in the small groups, while the nonauthoritarian man was less likely to participate in the permissive discussions.

A second study at Carleton College showed that students who think the instructor should be authoritarian tended to do poorly in independent study.

Although the present study utilizing the approach of method-individual interactions does not represent a major breakthrough, our ability tentatively to explain negative results coupled with the positive results we have so far achieved are encouraging. The teaching-learning process is a very complex one—one which can be mastered only through research designs permitting the study of teacher-group interactions as well as method-student interactions. Perhaps there is some virtue in the current college enrollment boom, for multivariate designs require *N*s much larger than any we have used to date.

Even with more definitive results, however, we would not conclude that students with certain types of personality should be excluded from independent study or small group discussions. As we see it, our goal should be for all students to learn to work independently and to participate responsibly in small groups. Rather than excluding students who dislike independence or work in small groups from these classes, we may want to give them special training and attention in order to help them learn how to learn in these situations. Increased knowledge about student personalities should give us increased ability to achieve these goals.

Received December 2, 1958.

[4] Jean Calloway directed the Carleton College study. Rae and Steven Kaplan carried out this analysis.

Psychology and Mental Health

A report of the Institute on Education and Training for Psychological Contributions to Mental Health, held at Stanford University in August, 1955.

Edited by CHARLES R. STROTHER

Topics discussed at the Institute:

Training Needs of Psychologists in Community Mental Health

Specialization in Training

Practicum Training

Training for Therapy

Training for Research in the Mental Health Field

Problems of Departmental Organization

Price: $1.75

Order from:

American Psychological Association, Inc.
1333 Sixteenth Street, N. W.
Washington 6, D. C.

Again Available

SCHOOL PSYCHOLOGISTS AT MID-CENTURY

Edited by
NORMA E. CUTTS

The Report of the Thayer Conference, published in 1955, covered an intensive study of the development of school psychology. Because of continuing demand, the Report has been reprinted.

230 pages. Price $2.75

AMERICAN PSYCHOLOGICAL ASSOCIATION
1333 Sixteenth Street, N.W.
Washington 6, D. C.

GRADUATE EDUCATION IN PSYCHOLOGY

Report of the Conference on Graduate Education in Psychology, sponsored by the Education and Training Board of the American Psychological Association and supported by a grant from the National Institute of Mental Health, U. S. Public Health Service; held at Miami Beach, Florida, November 29 to December 7, 1958.

Prepared by the Editorial Committee:
Anne Roe, *Chairman*,
and
John W. Gustad, Bruce V. Moore,
Sherman Ross, and Marie Skodak

Price $1.50

THE AMERICAN
PSYCHOLOGICAL ASSOCIATION
Dept. Grad
1333 Sixteenth Street, N.W.
Washington 6, D. C.

The
Journal
of
Educational
Psychology